OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: June 9, 2003, 08:36:50 ; Search time 474 Seconds
                 (without alignments)
                 1289.365 Million cell updates/sec

Title: US-09-550-163-1_COPY_230_250
Perfect score: 21
Sequence: 1 ctcatggtgatgattcggaatg 21

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 1455140287 8 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
          1: gb_ba:*
          2: gb_htg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*
         15: em_ba:*
         16: em_fun:*
         17: em_hum:*
         18: em_in:*
         19: em_mu:*
         20: em_om:*
         21: em_or:*
         22: em_ov:*
         23: em_pat:*
         24: em_ph:*
         25: em_pl:*
         26: em_ro:*
         27: em_sts:*
         28: em_un:*
         29: em_vi:*
         30: em_htg_hum:*
         31: em_htg_inv:*
         32: em_htg_other:*
         33: em_htg_mus:*
         34: em_htg_pln:*
         35: em_htg_rod:*
         36: em_htg_mam:*
         37: em_htg_vrt:*
         38: em_sy:*
         39: em_htgo_hum:*
         40: em_htgo_mus:*
         41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21	100.0	732	6	AX406939	AX406939 Sequence
2		21	100.0	732	6	AX406941	AX406941 Sequence
3		21	100.0	732	6	AX406947	AX406947 Sequence
4		21	100.0	732	6	AF071002	AF071002 Homo sapi
5		21	100.0	809	9	AF302095	AF302095 Homo sapi
6		21	100.0	24608	9	AP000320	AP000320 Homo sapi
7		21	100.0	100000	9	AP000052	AP000052 Homo sapi
8		21	100.0	100000	9	AP000167	AP000167 Homo sapi
9		21	100.0	100000	17	AP000120	AP000120 Homo sapi
10		21	100.0	340000	9	AP001719	AP001719 Homo sapi
11		19.4	92.4	215	4	AF329636	AF329636 Oryctolag
12		19.4	92.4	225	4	AY079211	AY079211 Sus scrof
13		19.4	92.4	732	6	AX406943	AX406943 Sequence
14		19.4	92.4	732	6	AX406945	AX406945 Sequence
15		19.4	92.4	117120	2	AC096260	AC096260 Rattus no
16	c	18.4	87.6	192351	5	AL772193	AL772193 Zebrafish
17		18.4	87.6	217073	2	AC094162	AC094162 Rattus no
18		18	85.7	4425	8	OSA427978	AJ427978 Oryza sat
19	c	18	85.7	91370	2	AP003757	AP003757 Oryza sat
20		18	85.7	229963	2	AC125174	AC125174 Mus muscu
21		17.8	84.8	372	10	AY050513	AY050513 Cavia por
22		17.8	84.8	468	10	AF071003	AF071003 Rattus no
23		17.8	84.8	1664	10	BC022699	BC022699 Mus muscu
24		17.8	84.8	2172	4	ECA292081	AJ292081 Equus cab
25		17.8	84.8	3883	5	AF428143	AF428143 Salmo sal
26		17.8	84.8	106925	2	AC116240	AC116240 Rattus no
27	c	17.8	84.8	133909	2	AC107430	AC107430 Rattus no
28		17.8	84.8	141726	2	AC110386	AC110386 Rattus no
29	c	17.8	84.8	144709	2	AC117904	AC117904 Rattus no
30	c	17.8	84.8	156140	9	AC067745	AC067745 Homo sapi
31	c	17.8	84.8	174073	2	AC107534	AC107534 Rattus no
32		17.8	84.8	186526	10	ALS96083	AL596083 Mouse DNA
33	c	17.4	82.9	215467	2	AC013420	AC013420 Homo sapi
34	c	17.4	82.9	218215	2	AC098542	AC098542 Rattus no
35	c	17.4	82.9	223496	2	AL732573	AL732573 Mus muscu
36		17.4	82.9	1302	14	AF222622	AF222622 Influenza
37		17.4	82.9	1302	14	AF222623	AF222623 Influenza
38		17.4	82.9	1302	14	AF222624	AF222624 Influenza
39		17.4	82.9	1302	14	AF222625	AF222625 Influenza
40		17.4	82.9	1302	14	AF222626	AF222626 Influenza
41		17.4	82.9	2027	14	AF046086	AF046086 Influenza
42		17.4	82.9	2089	14	AF098578	AF098578 Influenza
43		17.4	82.9	2250	14	AY043030	AY043030 Influenza
44		17.4	82.9	2253	14	AF156435	AF156435 Influenza
45		17.4	82.9	2253	14	AF156440	AF156440 Influenza

ALIGNMENTS

RESULT 1
AX406939
LOCUS   AX406939   732 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0222875.
ACCESSION   AX406939
VERSION   AX406939.1 GI:21439814
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Goldstein.S.A.
  TITLE   Polymorphisms associated with cardiac arrythmia
  JOURNAL   Patent: WO 0222875-A 1 21-MAR-2002;

Pred. No. is the number of results predicted by chance to have a

```
FEATURES        Location/Qualifiers
     source     1..732
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
     CDS        74..445
                /note="minK-related peptide 1, Val substituted for Ala at
                amino acid 116"
                /codon_start=1
                /protein_id="CAD35310.1"
                /db_xref="GI:21439815"
                /translation="MSTLSNFQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                NFYYVILYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                LEESKATIHENIGVAGFKMSP"
     variation  420
                /note="The drug associated here was quinidine."
                /replace="c"
BASE COUNT    221 a    151 c    157 g    203 t
ORIGIN
Query Match
Best Local Similarity   100.0%;   Score 21;  DB 6;  Length 732;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1 CTCATGGTGATGATTGGAATG 21
           |||||||||||||||||||||
Db     230 CTCATGGTGATGATTGGAATG 250

RESULT 2
AX406941
LOCUS       AX406941          732 bp    DNA          linear     PAT 14-JUN-2002
DEFINITION  Sequence 3 from Patent WO0222875.
ACCESSION   AX406941
VERSION     AX406941.1  GI:21439816
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Goldstein,S.A.
  TITLE     Polymorphisms associated with cardiac arrythmia
  JOURNAL   Patent: WO 0222875-A 3 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES        Location/Qualifiers
     source     1..732
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
     CDS        74..445
                /note="minK-related peptide 1, wild type"
                /codon_start=1
                /protein_id="CAD35180.1"
                /db_xref="GI:21439817"
                /translation="MSTLSNFQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                NFYYVILYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                LEESKATIHENIGAAGFKMSP"
BASE COUNT    221 a    152 c    158 g    202 t
ORIGIN
Query Match
Best Local Similarity   100.0%;   Score 21;  DB 6;  Length 732;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1 CTCATGGTGATGATTGGAATG 21
           |||||||||||||||||||||
Db     230 CTCATGGTGATGATTGGAATG 250

RESULT 3
AX406947
LOCUS       AX406947          732 bp    DNA          linear     PAT 14-JUN-2002
DEFINITION  Sequence 9 from Patent WO0222875.
```

```
ACCESSION   AX406947
VERSION     AX406947.1  GI:21439822
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Goldstein,S.A.
  TITLE     Polymorphisms associated with cardiac arrythmia
  JOURNAL   Patent: WO 0222875-A 9 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES        Location/Qualifiers
     source     1..732
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
     CDS        74..445
                /note="minK-related peptide 1, Ala substituted for Thr at
                amino acid 8"
                /codon_start=1
                /protein_id="CAD35312.1"
                /db_xref="GI:21439823"
                /translation="MSTLSNFAQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                NFYYVILYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                LEESKATIHENIGAAGFKMSP"
     variation  95
                /note="The drug associated here is was Bactrim."
                /replace="a"
BASE COUNT    220 a    152 c    158 g    202 t
ORIGIN
Query Match
Best Local Similarity   100.0%;   Score 21;  DB 6;  Length 732;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1 CTCATGGTGATGATTGGAATG 21
           |||||||||||||||||||||
Db     230 CTCATGGTGATGATTGGAATG 250

RESULT 4
AF071002
LOCUS       AF071002          732 bp    mRNA          linear     PRI 29-APR-1999
DEFINITION  Homo sapiens minK-related peptide 1 mRNA, complete cds.
ACCESSION   AF071002
VERSION     AF071002.1  GI:4704422
KEYWORDS    .
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 732)
  AUTHORS   Abbott,G.W., Sesti,F., Splawski,I., Buck,M.E., Lehmann,M.H.,
            Timothy,K.W., Keating,M.T. and Goldstein,S.A.
  TITLE     MiRP1 forms IKr potassium channels with HERG and is associated with
            cardiac arrythmia
  JOURNAL   Cell 97 (2), 175-187 (1999)
   MEDLINE  99235979
   PUBMED   10219239
REFERENCE   2  (bases 1 to 732)
  AUTHORS   Abbott,G.W., Sesti,F., Buck,M.E. and Goldstein,S.A.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1998) Section of Developmental Biology and
            Biophysics, Department of Pediatrics and Boyer Center for Molecular
            Medicine, Yale University School of Medicine, 295 Congress Avenue,
            New Haven, CT 06536, USA
FEATURES        Location/Qualifiers
     source     1..732
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="21"
                /map="21q11.1"
                /tissue_type="heart"
```

```
CDS                  74..445
                     /note="potassium channel subunit; MiRP1"
                     /codon_start=1
                     /product="minK-related peptide 1"
                     /protein_id="AAD28086.1"
                     /db_xref="GI:4704423"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYVIYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
BASE COUNT         221 a        152 c        157 g        202 t
ORIGIN

Query Match         100.0%;   Score 21;   DB 9;   Length 732;
Best Local Similarity  100.0%;   Pred. No. 5.8;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db        230 CTCATGGTGATGATTGGAATG 250


RESULT 5
AF302095
LOCUS            AF302095               809 bp    mRNA    linear   PRI 14-SEP-2000
DEFINITION       Homo sapiens voltage-gated K+ channel subunit MIRP1 (KCNE2) mRNA,
                 complete cds.
ACCESSION        AF302095
VERSION          AF302095.1  GI:10121887
KEYWORDS         .
SOURCE           Homo sapiens.
ORGANISM         Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1  (bases 1 to 809)
AUTHORS          Domenech,A., Estivill,X. and de la Luna,S.
TITLE            Cloning of human MIRP1 cDNA
JOURNAL          Unpublished
REFERENCE        2  (bases 1 to 809)
AUTHORS          Domenech,A., Estivill,X. and de la Luna,S.
TITLE            Direct Submission
JOURNAL          Submitted (01-SEP-2000) Medical and Molecular Genetics Center,
                 Institut Recerca Oncologica, Avia. de Castelldefels Km 2,7,
                 L'Hospitalet de Llobregat, Barcelona 08907, Spain
FEATURES             Location/Qualifiers
     source          1..809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22"
     gene            1..809
                     /gene="KCNE2"
     CDS             141..512
                     /gene="KCNE2"
                     /codon_start=1
                     /product="voltage-gated K+ channel subunit MIRP1"
                     /protein_id="AAG13416.1"
                     /db_xref="GI:10121888"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYVIYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
BASE COUNT         247 a        172 c        189 g        200 t      1 others
ORIGIN

Query Match         100.0%;   Score 21;   DB 9;   Length 809;
Best Local Similarity  100.0%;   Pred. No. 5.8;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db        297 CTCATGGTGATGATTGGAATG 317


RESULT 6
AP000320
LOCUS            AP000320               24608 bp    DNA    linear   PRI 20-NOV-1999
DEFINITION       Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
                 clone:Q12C8, complete sequence.
ACCESSION        AP000320
VERSION          AP000320.1  GI:4835689
KEYWORDS         HTG.
SOURCE           Homo sapiens DNA, clone:Q12C8.
ORGANISM         Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1  (bases 1 to 24608)
AUTHORS          Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
                 Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE            Homo sapiens 24,608bp genomic DNA of 21q22.1
JOURNAL          Published Only in DataBase (1999)
REFERENCE        2  (bases 1 to 24608)
AUTHORS          Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
                 Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE            Direct Submission
JOURNAL          Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
                 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
                 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
                 Japan (E-mail:hattori@gsc.riken.go.jp,
                 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
                 Fax:81-42-778-9924)
COMMENT          The sequence is a part of the data (ACCESSION No. AP000165 -
                 AP000173).
                 The sequencing project is supported by Japan Science Technology
                 Corporation (JST) and the Institute of Physical and Chemical
                 Research (RIKEN).
FEATURES             Location/Qualifiers
     source          1..24608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="Q12C8"
BASE COUNT        6797 a       5420 c       5437 g       6954 t
ORIGIN

Query Match         100.0%;   Score 21;   DB 9;   Length 24608;
Best Local Similarity  100.0%;   Pred. No. 7.1;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db      15852 CTCATGGTGATGATTGGAATG 15872


RESULT 7
AP000052
LOCUS            AP000052              100000 bp    DNA    linear   PRI 20-NOV-1999
DEFINITION       Homo sapiens genomic DNA, chromosome 21q22.1, segment 23/28,
                 complete sequence.
ACCESSION        AP000052
VERSION          AP000052.1  GI:3132362
KEYWORDS         HTG.
SOURCE           Homo sapiens DNA, clone:245P17-f4A4f_2.
ORGANISM         Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1  (bases 1 to 100000)
AUTHORS          Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE            Homo sapiens genomic DNA, chromosome 21q
JOURNAL          Published Only in DataBase (1998)
REFERENCE        2  (bases 1 to 100000)
AUTHORS          Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE            Direct Submission
JOURNAL          Submitted (11-MAY-1998) Masahira Hattori, Kitasato University,
                 Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
                 Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
                 Tel:0427-78-9732, Fax:0427-78-9561)
```

COMMENT   This sequence is conducted by Kitasato University JST sequencing
             Laboratory as a JST sequencing team.
             Principal Investigator:Yoshiyuki Sakaki Ph.D.
             Phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
             sakaki@hgc.ims.u-tokyo.ac.jp
             Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The
             sequence is submitted by:Human Genome Sequencing in ALIS project of
             JST
             Japan Science and Technology Corporation (JST)
             5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan
             For further information about this sequence, including its location
             and relationship to other sequences, please visit our sequence
             archive Web site (http://www.alis.tokyo.jst.go.jp/HGS/top.html)
             or send email to webmaster@www-alis.tokyo.jst.go.jp'.

FEATURES   Location/Qualifiers
     source 1..100000
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="21"
             /map="21q22.1"
             /clone="245P17-F4A4f_2"

BASE COUNT   27603 a   21934 c   22513 g   27950 t
ORIGIN

Query Match   100.0%;   Score 21; DB 9; Length 100000;
Best Local Similarity   100.0%;   Pred. No. 7.7;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
        |||||||||||||||||||||
Db 80374 CTCATGGTGATGATTGGAATG 80394

RESULT 8
AP000167
LOCUS   AP000167   100000 bp   DNA   linear   PRI 08-JAN-2000
DEFINITION Homo sapiens genomic DNA, chromonome 21q22.1, D21S226-AML region,
            clone B2344F14-f50E8, segment 3/9, complete sequence.
ACCESSION   AP000167
VERSION   AP000167.1   GI:4827132
KEYWORDS   HTG.
SOURCE   Homo sapiens DNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 100000)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE   Homo sapiens 890,291bp genomic DNA of 21q22.1 (REGION: D21S226-AML
            CLONE RANGE: B2344F14-f50E8)
JOURNAL   Published Only in DataBase (1999)
REFERENCE   2 (bases 1 to 100000)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Ginomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT   E. coli transposon insertion:The present data does not contain E.
            coli transposon sequences which integrated in the
            original/previous sequences. We determined the boundary between
            the insertion and genomic sequences experimentally, removed the
            insertion sequences, reconstitutd the present data. The sequencing
            project is supported by Japan Science Technology Corporation (JST)
            and The Institute of Physical and Chemical Research (RIKEN).

FEATURES   Location/Qualifiers
     source 1..100000
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="21"

BASE COUNT   27338 a   22306 c   22571 g   27785 t
ORIGIN   /map="21q22.1"

Query Match   100.0%;   Score 21; DB 9; Length 100000;
Best Local Similarity   100.0%;   Pred. No. 7.7;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
        |||||||||||||||||||||
Db 72342 CTCATGGTGATGATTGGAATG 72362

RESULT 9
AP000120
ID   AP000120   standard; DNA; HUM; 100000 BP.
XX
AC   AP000120;
XX
SV   AP000120.1
XX
DT   04-MAY-1999 (Rel. 59, Created)
DT   26-SEP-1999 (Rel. 61, Last updated, Version 3)
XX
DE   Homo sapiens genomic DNA of 21q22.1, GART and AML related, SLC5A3-f4A4
DE   region, segment 3/8.
XX
KW   HTG.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-100000
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RT   ;
RL   Submitted (15-APR-1999) to the EMBL/GenBank/DDBJ databases.
RL   Mika Hirakawa, Japan Science and Technology Corporation (JST), Advanced
RL   Databases Department; 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081, Japan
RL   (E-mail:mika@tokyo.jst.go.jp, URL:http://www-alis.tokyo.jst.go.jp/,
RL   Tel:81-3-5214-8491, Fax:81-3-5214-8470)
XX
RN   [2]
RA   Hattori M., Ishii K., Toyoda A., Shiba T., Sakaki Y.;
RT   "Homo sapiens 817,199bp genomic DNA of 21q22.1 GART and AML region";
RL   Unpublished.
XX
CC   This sequence is conducted by Kitasato University JST sequencing
CC   Laboratory as a JST sequencing team.
CC   Principal Investigator:Yoshiyuki Sakaki Ph.D.
CC   Phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
CC   sakaki@hgc.ims.u-tokyo.ac.jp
CC   Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D.
CC   The sequence is submitted by Human Genome Sequencing in ALIS
CC   project of JST.
CC   Japan Science and Technology Corporation (JST)
CC   5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081 Japan
CC   For further information about this sequence, including its
CC   location and relationship to other sequences, please visit our
CC   sequence archive Web site (http://www.alis.tokyo.jst.go.jp/HGS/)
CC   or send email to webmaster@www-alis.tokyo.jst.go.jp
XX
FH   Key   Location/Qualifiers
FH
FT   source   1..100000
FT   /chromosome="21"
FT   /db_xref="taxon:9606"
FT   /sequenced_mol="DNA"
FT   /organism="Homo sapiens"
FT   /map="21q22.1"
FT   STS   32213..32320
FT   /note="83F9aL-F/83F9aL-R;The location is between each

```
FT                flanking site of PCR primers."
FT     STS        /standard_name="D21S1689"
FT                43934..44157
FT                /note="AFMa08eyf9;Genethon Marker;The location is between
FT                each flanking site of PCR primers."
FT                /standard_name="D21S1920"
XX
SQ     Sequence 100000 BP; 27731 A; 21990 C; 22220 G; 28059 T; 0 other;

Query Match     100.0%;   Score 21;  DB 17;  Length 100000;
Best Local Similarity 100.0%;  Pred. No. 7.7;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 CTCATGTGATGATTGGAATG 21
            ||||||||||||||||||||
Db  89676 CTCATGTGATGATTGGAATG 89696


RESULT 10
AP001719
LOCUS       AP001719               340000 bp    DNA             linear   PRI 10-MAY-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 21q, section 63/105.
ACCESSION   AP001719 AL163264 BA000005
VERSION     AP001719.1  GI:7768719
KEYWORDS    .
SOURCE      Homo sapiens DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
AUTHORS     Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt.P., Scharfe,M., Schoen.O., Desario,A.,
            Reichelt,J., Kauer.G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., and
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
TITLE       The DNA sequence of human chromosome 21
JOURNAL     Nature 405 (6784), 311-319 (2000)
MEDLINE     20289799
REFERENCE   2  (bases 1 to 340000)
AUTHORS     Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K., and
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
TITLE       Direct Submission
JOURNAL     Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing
            Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
            Group * Institute of Molecular Biotechnology, Genome Analysis *
            Keio University School of Medicine, Dept. of Molecular Biology *
            GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
            Genetics (addresses see below)
COMMENT     On May 30, 2000 this sequence version replaced gi:7717333.
            The chromosome 21 mapping and sequencing consortium consisting of
            * RIKEN Genomic Sciences Center, Human Genome Research Group, *
            Sagamihara 228-8555, Japan,
            * e.mail: hattori@gsc.riken.go.jp
            * URL: http://hgp.gsc.riken.go.jp/
```

```
            and
            * Institute of Molecular Biotechnology, Genome Analysis, *
            Beutenbergstrasse 11, D-07745 Jena, Germany,
            * e.mail: gscj-submit@genome.imb-jena.de
            * URL: http://genome.imb-jena.de/
            and
            * Keio University School of Medicine, Molecular Biology, * Tokyo
            160-8582, Japan,
            * e.mail: nshimizu@dmb-med.keio.ac.jp
            * URL: http://www.dmb.med.keio.ac.jp/
            and
            * GBF, Dept. of Genome Analysis,
            * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
            info.genome@gbf.de
            * URL: http://genome.gbf.de/
            and
            * Max-Planck Institute for Molecular Genetics,
            * Ihnestrasse 73, D-14195 Berlin, Germany,
            * e.mail: info-chr21@molgen.mpg.de
            * URL: http://chr21.rz-berlin.mpg.de/
AL163264: Submitted (10-Apr-2000).
FEATURES             Location/Qualifiers
     source          1..340000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     source          <1..29420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R338L7, 5' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000467"
     source          29242..97350
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="CTD2344F14"
                     /clone_lib="CTD BAC library"
                     /note="Accession No. AP000315"
     source          96883..121400
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pS185"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000316"
     source          121210..218631
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="P245P17"
                     /clone_lib="RP1 PAC library"
                     /note="Accession No. AP000317"
     source          218558..260079
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ6F8"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000318"
     source          258735..285904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ94A10"
                     /clone_lib="CMP21 P1 library"
```

```
                                        /rpt_type=TANDEM


                Query Match              100.0%;  Score 21;  DB 9;  Length 340000;
                Best Local Similarity    100.0%;  Pred. No. 8.2;
                Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

                Qy        1 CTCATGGTGATGATTGGAATG 21
                            |||||||||||||||||||||
                Db   301583 CTCATGGTGATGATTGGAATG 301603

RESULT 11
AF329636
LOCUS       AF329636                 215 bp    mRNA    linear   MAM 28-SEP-2001
DEFINITION  Oryctolagus cuniculus K+/pacemaker channel beta subunit mirp1
            (kcne2) mRNA, partial cds.
ACCESSION   AF329636
VERSION     AF329636.1  GI:13194729
KEYWORDS    .
SOURCE      Oryctolagus cuniculus.
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Yu,H., Wu,J., Potapova,I., Wymore,R.T., Holmes,B., Zuckerman,J.,
            Pan,Z., Wang,H., Shi,W., Robinson,R.B., El-Maghrabi,M.R.,
            Benjamin,W., Dixon,J., McKinnon,D., Cohen,I.S. and Wymore,R.S.
  TITLE     MinK-related peptide 1: A beta subunit for the HCN ion channel
            subunit family enhances expression and speeds activation
  JOURNAL   Circ. Res. 88 (12), E84-E87 (2001)
  MEDLINE   21313430
  PUBMED    11420311
REFERENCE   2  (bases 1 to 215)
  AUTHORS   Wymore,R.T., Holmes,B.A., Wymore,R.S., Yu,H., Wu,J., Potapova,I.,
            Zuckerman,J., Pan,Z., Wang,H., Shi,W., Robinson,R., El-Maghrabi,R.,
            Benjamin,W., Dixon,J.E., McKinnon,D. and Cohen,I.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2000) Biology, The University of Tulsa, 600 S.
            College Av., Tulsa, OK 74104-3189, USA
FEATURES             Location/Qualifiers
     source          1..215
                     /organism="Oryctolagus cuniculus"
                     /db_xref="taxon:9986"
     gene            <1..>215
                     /gene="kcne2"
     CDS             <1..>215
                     /gene="kcne2"
                     /codon_start=3
                     /product="K+/pacemaker channel beta subunit mirp1"
                     /protein_id="AAK15527.1"
                     /db_xref="GI:13194730"
                     /translation="AENFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQ
                     YIVEDWQEKYKSQILHFEEAKATIHEN"
BASE COUNT       60 a      58 c     51 g     46 t
ORIGIN

                Query Match              92.4%;  Score 19.4;  DB 4;  Length 215;
                Best Local Similarity    95.2%;  Pred. No. 37;
                Matches  20;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

                Qy        1 CTCATGGTGATGATTGGAATG 21
                            |||||||||||||||| ||||
                Db       33 CTCATGGTGATGATTGGCATG 53

RESULT 12
AY079211
LOCUS       AY079211                 225 bp    mRNA    linear   MAM 20-JUL-2002
DEFINITION  Sus scrofa KCNE2 mRNA, partial cds.
ACCESSION   AY079211
VERSION     AY079211.1  GI:21913149
KEYWORDS    .
SOURCE      pig.
```

```
source         /note="Accession No. AP000319"
               285732..310339
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="21"
               /map="21q22.1"
               /clone="pQ12C8"
               /clone_lib="CMP21 P1 library"
               /note="Accession No. AP000320"
source         310050..326194
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="21"
               /map="21q22.1"
               /clone="pQ82F5"
               /clone_lib="CMP21 P1 library"
               /note="Accession No. AP000321"
source         325966..>340000
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="21"
               /map="21q22.1"
               /clone="pQ97G8, 3' partial"
               /clone_lib="CMP21 P1 library"
               /note="Accession No. AP000322"
repeat_region  complement(1182..1688)
               /note="AluSp"
               /rpt_family="SINE/Alu"
               /rpt_type=DISPERSED
repeat_region  complement(2099..2390)
               /note="AluSx"
               /rpt_family="SINE/Alu"
               /rpt_type=DISPERSED
repeat_region  2799..2884
               /note="MIR"
               /rpt_family="SINE/MIR"
               /rpt_type=DISPERSED
repeat_region  3288..3381
               /note="MIR"
               /rpt_family="SINE/MIR"
               /rpt_type=DISPERSED
repeat_region  complement(2891..3287)
               /note="MER31A"
               /rpt_family="LTR/MER4-group"
               /rpt_type=DISPERSED
misc_feature   4211..5231
               /note="CpG island"
misc_feature   4272..4279
               /note="NotI site"
repeat_region  4328..4476
               /note="(CGGGG)n"
               /rpt_family="Simple_repeat"
               /rpt_type=TANDEM
gene           4392..35127
               /gene="SLC5A3"
mRNA           join(4392..4740,12254..35127)
               /gene="SLC5A3"
               /product="Sodium/myo-inositol cotransporter"
               /note="human solute carrier family 5, member 3, Accession
               No. AF027153"
exon           4392..4740
               /gene="SLC5A3"
               /number=1
misc_feature   4392..4399
               /gene="SLC5A3"
               /note="NotI site"
repeat_region  4810..4985
               /note="(CGGGG)n"
               /rpt_family="Simple_repeat"
               /rpt_type=TANDEM
repeat_region  5099..5191
               /note="(CCG)n"
               /rpt_family="Simple_repeat"
```

```
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 225)
AUTHORS     Li,Y., Mitchell,K.E., Wymore,R.S. and Freeman,L.C.
TITLE       Granulosa cells express multiple KCNQ and KCNE channel subunits
            (Abstract 1220)
JOURNAL     Biophys. J. 82, 252a (2002)
REFERENCE   2  (bases 1 to 225)
AUTHORS     Li,Y., Wymore,T., Mitchell,K.E., Wymore,R.S. and Freeman,L.C.
TITLE       Direct Submission
JOURNAL     Submitted (19-FEB-2002) Anatomy and Physiology, College of
            Veterinary Medicine, Kansas State University, 228 Coles Hall, 1600
            Denison Ave., Manhattan, KS 66506-5802, USA
FEATURES    Location/Qualifiers
    source  1..225
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
    CDS     <1..>225
            /codon_start=1
            /product="KCNE2"
            /protein_id="AAM77423.1"
            /db_xref="GI:21913150"
            /translation="DAENFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYH
            QYIVEDWQGKYKSQILNLEESRATIHENLGA"
BASE COUNT  66 a   60 c   54 g   45 t
ORIGIN

Query Match             92.4%;  Score 19.4;  DB 4;  Length 225;
  Best Local Similarity 95.2%;  Pred. No. 37;
  Matches 20;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              ||||| |||||||||||||||
Db         34 CTGATGGTGATGATTGGAATG 54


RESULT 13
LOCUS       AX406943                 732 bp    DNA       linear   PAT 14-JUN-2002
DEFINITION  Sequence 5 from Patent WO0222875.
ACCESSION   AX406943
VERSION     AX406943.1  GI:21439818
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 5 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
    source  1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
    CDS     74..445
            /note="minK-related peptide 1, Thr substitued for Met at
            amino acid 54"
            /codon_start=1
            /protein_id="CAD35181.1"
            /db_xref="GI:21439819"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
            NFYYVILYLFVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
            LEESKATIHENIGAAGFKMSP"
    variation   234
            /note="The drug associated here was procainamide."
            /replace="t"
BASE COUNT  221 a   153 c   157 g   201 t
ORIGIN

Query Match             92.4%;  Score 19.4;  DB 6;  Length 732;
  Best Local Similarity 95.2%;  Pred. No. 39;
  Matches 20;  Conservative 0;  Mismatches 1;  Indels 1;  Gaps 0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              ||||| ||||||||||||||
Db        230 CTCACGGTGATGATTGGAATG 250


RESULT 14
LOCUS       AX406945                 732 bp    DNA       linear   PAT 14-JUN-2002
DEFINITION  Sequence 7 from Patent WO0222875.
ACCESSION   AX406945
VERSION     AX406945.1  GI:21439820
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 7 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
    source  1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
    CDS     74..445
            /note="minK-related peptide 1, Thr substituted for Ile at
            amino acid 57"
            /codon_start=1
            /protein_id="CAD35311.1"
            /db_xref="GI:21439821"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
            NFYYVILYLMVMTGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
            LEESKATIHENIGAAGFKMSP"
    variation   243
            /note="The drug associated here was oxatomide."
            /replace="t"
BASE COUNT  221 a   153 c   157 g   201 t
ORIGIN

Query Match             92.4%;  Score 19.4;  DB 6;  Length 732;
  Best Local Similarity 95.2%;  Pred. No. 39;
  Matches 20;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy          1 CTCATGGTGATGATTGGAATG 21
              ||||||||||||| ||||||
Db        230 CTCATGGTGATGACTGGAATG 250


RESULT 15
LOCUS       AC096260              117120 bp    DNA       linear   HTG 11-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-78O12, *** SEQUENCING IN PROGRESS
            ***, 60 unordered pieces.
ACCESSION   AC096260
VERSION     AC096260.3  GI:21723414
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 117120)
AUTHORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,S., Brown,M., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
```

TITLE

AUTHORS   Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
          Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
          Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
          Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
          Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
          Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
          Gabisi,A., Gao,J., Garcia,A., Gainer,T., Garza,N., Gill,R.,
          Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
          Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
          Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
          Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
          Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
          Karlsson,B., Kelly,S., Khan,U., King,U., Korvah,J., Kovar,C.,
          Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
          Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
          Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
          Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
          Massey,E., Mawhiney,S., McLeod,M.P., Meador,M.P., Mei,G., Metzker,M.,
          Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
          Moser,M., Neal,D., Newton,J., Nwtson,N., Nguyen,A., Nguyen,N.,
          Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
          Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
          Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
          Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
          Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,D.,
          Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
          Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
          Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
          Williams,G., Williamson,A., Wlecyk,R., Wooden,S., Worley,K.,
          Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
          Weinstock,G. and Gibbs,R.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 117120)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 117120)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA

COMMENT   On Jul 10, 2002 this sequence version replaced gi:17943956.
          ---------------------------- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc-help@bcm.tmc.edu
          ---------------------------- Project Information
          Center project name: GEUJ
          Center clone name: CH230-78O12
          ---------------------------- Summary Statistics
          Sequencing vector: Plasmid;
          Chemistry: Dye-terminator Big Dye: 100% of reads
          Assembly program: Phrap; version 0.990329
          Consensus quality: 69897 bases at least Q40
          Consensus quality: 76882 bases at least Q30
          Consensus quality: 81814 bases at least Q20
          ----
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 60 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will

```
* be preserved.
*           1   1481:  contig of 1481 bp in length
*        1482   1581:  gap of unknown length
*        1582   2737:  contig of 1156 bp in length
*        2738   2837:  gap of unknown length
*        2838   4522:  contig of 1685 bp in length
*        4523   4622:  gap of unknown length
*        4623   5803:  contig of 1181 bp in length
*        5804   5903:  gap of unknown length
*        5904   6978:  contig of 1075 bp in length
*        6979   7078:  gap of unknown length
*        7079   8789:  contig of 1711 bp in length
*        8790   8889:  gap of unknown length
*        8890  10439:  contig of 1550 bp in length
*       10440  10539:  gap of unknown length
*       10540  11546:  contig of 1007 bp in length
*       11547  11646:  gap of unknown length
*       11647  12988:  contig of 1342 bp in length
*       12989  13088:  gap of unknown length
*       13089  14603:  contig of 1515 bp in length
*       14604  14703:  gap of unknown length
*       14704  15786:  contig of 1083 bp in length
*       15787  15886:  gap of unknown length
*       15887  17554:  contig of 1668 bp in length
*       17555  17654:  gap of unknown length
*       17655  18735:  contig of 1081 bp in length
*       18736  18835:  gap of unknown length
*       18836  20875:  contig of 2040 bp in length
*       20876  20975:  gap of unknown length
*       20976  22068:  contig of 1093 bp in length
*       22069  22168:  gap of unknown length
*       22169  23351:  contig of 1183 bp in length
*       23352  23451:  gap of unknown length
*       23452  24974:  contig of 1523 bp in length
*       24975  25074:  gap of unknown length
*       25075  26111:  contig of 1037 bp in length
*       26112  26211:  gap of unknown length
*       26212  27583:  contig of 1372 bp in length
*       27584  27683:  gap of unknown length
*       27684  29195:  contig of 1512 bp in length
*       29196  29295:  gap of unknown length
*       29296  30486:  contig of 1191 bp in length
*       30487  30586:  gap of unknown length
*       30587  32004:  contig of 1418 bp in length
*       32005  32104:  gap of unknown length
*       32105  33847:  contig of 1743 bp in length
*       33848  33947:  gap of unknown length
*       33948  35016:  contig of 1069 bp in length
*       35017  35116:  gap of unknown length
*       35117  36608:  contig of 1492 bp in length
*       36609  36708:  gap of unknown length
*       36709  38612:  contig of 1904 bp in length
*       38613  38712:  gap of unknown length
*       38713  39836:  contig of 1124 bp in length
*       39837  39936:  gap of unknown length
*       39937  41024:  contig of 1088 bp in length
*       41025  41124:  gap of unknown length
*       41125  42555:  contig of 1431 bp in length
*       42556  42655:  gap of unknown length
*       42656  44858:  contig of 2203 bp in length
*       44859  44958:  gap of unknown length
*       44959  46558:  contig of 1600 bp in length
*       46559  46658:  gap of unknown length
*       46659  48692:  contig of 2034 bp in length
*       48693  48792:  gap of unknown length
*       48793  50888:  contig of 2096 bp in length
*       50889  50988:  gap of unknown length
*       50989  52347:  contig of 1359 bp in length
*       52348  52447:  gap of unknown length
*       52448  54571:  contig of 2124 bp in length
*       54572  54671:  gap of unknown length
*       54672  56188:  contig of 1517 bp in length
*       56189  56288:  gap of unknown length
```

```
56289 *  57934: contig of 1646 bp in length
57935 *  58034: gap of unknown length
58035 *  59723: contig of 1689 bp in length
59724 *  59823: gap of unknown length
59824 *  61614: contig of 1791 bp in length
61615 *  61714: gap of unknown length
61715 *  63009: contig of 1295 bp in length
63010 *  63109: gap of unknown length
63110 *  64743: contig of 1634 bp in length
64744 *  64843: gap of unknown length
64844 *  66689: contig of 1846 bp in length
66690 *  66789: gap of unknown length
66790 *  68797: contig of 2008 bp in length
68798 *  68897: gap of unknown length
68898 *  70914: contig of 2017 bp in length
70915 *  71014: gap of unknown length
71015 *  73395: contig of 2381 bp in length
73396 *  73495: gap of unknown length
73496 *  75642: contig of 2147 bp in length
75643 *  75742: gap of unknown length
75743 *  78555: contig of 2813 bp in length
78556 *  78655: gap of unknown length
78656 *  80280: contig of 1625 bp in length
80281 *  80380: gap of unknown length
80381 *  83171: contig of 2791 bp in length
83172 *  83271: gap of unknown length
83272 *  85608: contig of 2337 bp in length
85609 *  85708: gap of unknown length
85709 *  87924: contig of 2216 bp in length
87925 *  88024: gap of unknown length
88025 *  90323: contig of 2299 bp in length
90324 *  90423: gap of unknown length
90424 *  93281: contig of 2858 bp in length
93282 *  93381: gap of unknown length
93382 *  96276: contig of 2895 bp in length
```

```
Query Match      92.4%;   Score 19.4;   DB 2;   Length 117120;
Best Local Similarity   95.2%;   Pred. No. 53;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy        1 CTCATGGTGATGATGATTGGAATG 21
            |||||||||||||||||| |||||
Db    70138 CTCATGGTGATGATGATTTGAATG 70158
```

Search completed: June 9, 2003, 09:12:03
Job time : 605 secs

OM nucleic - nucleic search, using sw model

```
          GenCore version 5.1.6
       Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: June   9, 2003, 08:36:49 ; Search time 119 Seconds
                          (without alignments)
                  397.411 Million cell updates/sec

```
Title:          US-09-550-163-1_COPY_230_250
Perfect score:  21
Sequence:       1 ctcatggtgatgattggaatg 21
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       2185239 seqs, 112599159 residues
```

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_101002:*
    1: /SIDS27/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
    2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
    3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
    4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
    5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
    6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
    7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
    8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
    9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
   10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
   11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
   12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
   13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
   14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
   15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
   16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
   17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
   18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
   19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
   20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
   21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
   22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
   23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
   24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB Length	DB	ID	Description
1	21	100.0	21	312	22	ABA49938	Human breast cell
2	21	100.0	21	312	22	ABA67856	Human foetal liver
3	21	100.0	21	312	22	ABA34913	Probe #11379 for g
4	21	100.0	21	312	22	AAK16264	Human brain expres
5	21	100.0	21	312	22	AAK42008	Human bone marrow
6	21	100.0	21	312	22	AAI22773	Probe #12706 for g
7	21	100.0	21	312	22	AAI48075	Probe #16761 used
8	21	100.0	21	312	22	AAI08446	Probe #8437 used t
9	21	100.0	21	312	24	ABS16039	Human genome-deriv
10	21	100.0	21	372	22	AAI24432	Probe #14365 for g
11	21	100.0	21	372	22	AAI09965	Probe #9956 used t
12	21	100.0	21	372	22	AAS00245	Human potassium ch
13	21	100.0	21	450	22	ABA44797	Human breast cell
14	21	100.0	21	450	22	ABA55252	Human foetal liver
15	21	100.0	21	450	22	ABA24997	Probe #3463 for ge
16	21	100.0	21	450	22	AAK03508	Human brain expres
17	21	100.0	21	450	22	AAK28962	Human bone marrow
18	21	100.0	21	450	22	AAI13549	Probe #3482 for ge
19	21	100.0	21	450	22	AAI34911	Probe #3597 used t
20	21	100.0	21	450	22	AAI03438	Probe #3429 used t
21	21	100.0	21	450	24	ABS03496	Human genome-deriv
22	21	100.0	21	471	22	AAI15256	Probe #5189 for ge
23	21	100.0	21	471	22	AAI04990	Probe #4981 used t
24	21	100.0	21	471	22	AAF80269	Nucleotide sequenc
25	21	100.0	21	600	22	ABA09192	Human MiRP1 homolo
26	21	100.0	21	600	22	AAK52645	Human polynucleoti
27	21	100.0	21	655	22	AAK51661	Human polynucleoti
28	21	100.0	21	732	21	AAC64071	Human potassium ch
29	21	100.0	21	732	21	AAC64083	Human potassium ch
30	21	100.0	21	732	21	AAC64086	Human potassium ch
31	21	100.0	21	732	24	ABK86573	cDNA encoding huma / Human potassium ch
32	21	100.0	21	732	24	AAD35169	Human KCNE2 mutant
33	21	100.0	21	732	24	AAD35170	Human KCNE2 wild t
34	21	100.0	21	732	24	AAD35173	Human KCNE2 mutant
35	19.4	92.4	21	732	21	AAC64084	Human potassium ch
36	19.4	92.4	21	732	24	AAC64072	Human potassium ch
37	19.4	92.4	21	732	24	AAD35171	Human KCNE2 mutant
38	19.4	90.5	21	732	24	AAD35172	Human KCNE2 mutant
39	19	90.5	19	20	22	AAF80272	Primer used to amp
40	17.8	84.8		65	24	ABN29241	Rat spliced transc
41	17.8	84.8		372	21	AAS00246	Rat potassium chan
42	17.8	84.8		468	21	AAC64072	Rat potassium chan
43	16.8	80.0		583	22	ABA50512	Human breast cell
44	16.8	80.0		583	22	ABA68463	Human foetal liver
45	16.8	80.0		583	22	ABA35452	Probe #13918 for g

ALIGNMENTS

```
RESULT 1
ABA49938
ID   ABA49938 standard; DNA; 312 BP.
XX
AC   ABA49938;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #8633.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO20015771-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
```

```
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes -
XX
PS   Claim 4; SEQ ID NO 8633; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             100.0%;  Score 21;  DB 22;  Length 312;
Best Local Similarity   100.0%;  Pred. No. 1.9;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 CTCATGGTGATGATTGGAATG 21
         |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 2
ABA67856
ID   ABA67856 standard; DNA; 312 BP.
XX
AC   ABA67856;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16161.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,   Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-483447/52.
XX

PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
PS   Claim 4; SEQ ID NO 16161; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             100.0%;  Score 21;  DB 22;  Length 312;
Best Local Similarity   100.0%;  Pred. No. 1.9;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 CTCATGGTGATGATTGGAATG 21
         |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 3
ABA34913
ID   ABA34913 standard; DNA; 312 BP.
XX
AC   ABA34913;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #13379 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,   Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 13379; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
```

CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.inc/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
   Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   104 CTCATGGTGTGATGATTGGAATG 124

RESULT 4
AAK16264
ID   AAK16264 standard; DNA; 312 BP.
XX
AC   AAK16264;
XX
DT   05-NOV-2001 (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16255.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 16255; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
   Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   104 CTCATGGTGTGATGATTGGAATG 124

DE   Human bone marrow expressed single exon probe SEQ ID NO: 16565.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 16565; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
   Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   104 CTCATGGTGTGATGATTGGAATG 124

RESULT 5
AAK42008
ID   AAK42008 standard; DNA; 312 BP.
XX
AC   AAK42008;
XX
DT   06-NOV-2001 (first entry)

RESULT 6
AAI22773
ID   AAI22773 standard; DNA; 312 BP.
XX
AC   AAI22773;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #12706 for gene expression analysis in human cervical cell sample.
XX

KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO20157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 12706; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches 21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 7
AAI48075
ID   AAI48075 standard; DNA; 312 BP.
XX
AC   AAI48075;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #16761 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO20157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX

PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 16761; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches 21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 8
AAI08446
ID   AAI08446 standard; DNA; 312 BP.
XX
AC   AAI08446;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #8417 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO20157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX

XX
PS   Claim 25; SEQ ID No 8437; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 9
ABS16039
ID   ABS16039 standard; DNA; 312 BP.
XX
AC   ABS16039;
XX
DT   19-AUG-2002 (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 16030.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180112P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX

XX
PS   Claim 4; SEQ ID No 16030; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe open reading frame of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   100.0%; Score 21; DB 24; Length 312;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 10
AAI24432
ID   AAI24432 standard; DNA; 372 BP.
XX
AC   AAI24432;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #14365 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX

PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 14365; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 372;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   157 CTCATGGTGATGATTGGAATG 177

RESULT 11
AAI09965
ID   AAI09965 standard; DNA; 372 BP.
XX
AC   AAI09965;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #9956 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024253.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.

PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 9956; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 372;
Best Local Similarity   100.0%; Pred. No. 1.9;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   157 CTCATGGTGATGATTGGAATG 177

RESULT 12
AAS00245
ID   AAS00245 standard; DNA; 372 BP.
XX
AC   AAS00245;
XX
DT   10-MAY-2001 (first entry)
XX
DE   Human potassium channel regulatory protein, Mink2, DNA sequence.
XX
KW   Human; Mink2; potassium channel; cardiac arrhythmia; hypertension; ds;
KW   angina; asthma; diabetes; renal insufficiency; urinary incontinence;
KW   irritable colon; epilepsy; cerebrovascular ischaemia; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..372
FT   /tag= a
FT   /product= "MINK2 potassium channel protein"
XX
PN   WO200114403-A1.
XX
PD   01-MAR-2001.
XX
PF   18-AUG-2000; 2000WO-US22799.
XX
PR   20-AUG-1999; 99US-0379201.
XX
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI   Ficker E, Wible B, Brown AM;
XX
DR   WPI; 2001-218424/22.
DR   P-PSDB; AAU00215.
XX
PT   Novel potassium channel gene termed MinK2 encoding potassium channel
PT   regulatory protein, useful for screening compounds that are useful for

PT   treating diseases caused by aberrant potassium activity -
XX
PS   Claim 1; Fig 9; 39pp; English.
XX
CC   The sequence represents the coding sequence of human potassium channel
CC   regulatory protein, Mink2. Mink2 sequence is useful for producing a
CC   potassium channel regulatory protein useful for in vitro or in vivo
CC   screening of agonistic or antagonistic compounds that are useful for
CC   treating diseases caused by aberrant potassium activity, such as human
CC   cardiac arrhythmias, hypertension, angina, asthma, diabetes, renal
CC   insufficiency, urinary incontinence, irritable colon, epilepsy,
CC   cerebrovascular ischaemia, and autoimmune disease.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%;   Score 21; DB 22; Length 372;
Best Local Similarity 100.0%;   Pred. No. 1.9;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
                |||||||||||||||||||||
Db   157 CTCATGGTGATGATTGGAATG 177

RESULT 13
ABA44797
ID   ABA44797 standard; DNA; 450 BP.
XX
AC   ABA44797;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3492.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes -
XX
PS   Claim 1; SEQ ID NO 3492; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for

CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match   100.0%;   Score 21; DB 22; Length 450;
Best Local Similarity 100.0%;   Pred. No. 2;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
                |||||||||||||||||||||
Db   312 CTCATGGTGATGATTGGAATG 332

RESULT 14
ABA55252
ID   ABA55252 standard; DNA; 450 BP.
XX
AC   ABA55252;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3557.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 3557; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Search completed: June   9, 2003, 08:51:25
Job time : 120 secs

Query Match   100.0%;   Score 21; DB 22; Length 450;
Best Local Similarity 100.0%;   Pred. No. 2;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
          ||||||||||||||||||||||
Db   312 CTCATGGTGATGATTGGAATG 332

RESULT 15
ABA24997
ID   ABA24997 standard; DNA; 450 BP.
XX
AC   ABA24997;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Probe #3463 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 3463; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match   100.0%;   Score 21; DB 22; Length 450;
Best Local Similarity 100.0%;   Pred. No. 2;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
          ||||||||||||||||||||||
Db   312 CTCATGGTGATGATTGGAATG 332

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:51 ; Search time 31.5 Seconds
                                      (without alignments)
                                      204.451 Million cell updates/sec

Title: US-09-550-163-1_COPY_230_250
Perfect score: 21
Sequence: 1 ctcatcggtgatgattggaatg 21

Scoring table: IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents_NA:*
             1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
             6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match	Length	DB	ID	Description
c 1	16.4	78.1		25002	4	US-08-961-527-48	Sequence 48, Appl
c 2	16.2	77.1		1141	2	US-08-966-316-3	Sequence 3, Appli
c 3	16.2	77.1		9495	1	US-08-27-829-1	Sequence 1, Appli
c 4	16.2	77.1		9495	5	PCT-US93-01544-1	Sequence 1, Appli
c 5	16.2	77.1		19718	4	US-08-961-527-99	Sequence 99, Appl
c 6	15.8	75.2		836	2	US-08-858-207A-180	Sequence 180, App
c 7	15.8	75.2		1569	2	US-08-923-772-1	Sequence 1, Appli
c 8	15.8	75.2		1569	4	US-09-385-287-1	Sequence 1, Appli
c 9	15.8	75.2		2013	4	US-09-423-890-3	Sequence 3, Appli
c 10	15.8	75.2		2465	4	US-09-423-890-9	Sequence 9, Appli
c 11	15.8	75.2		2503	1	US-08-472-934-3	Sequence 3, Appli
c 12	15.8	75.2		2503	1	US-08-472-934-11	Sequence 11, Appl
c 13	15.8	75.2		2503	2	US-08-323-460A-3	Sequence 3, Appli
c 14	15.8	75.2		2503	2	US-08-461-146C-3	Sequence 3, Appli
c 15	15.8	75.2		2503	3	US-08-461-146C-11	Sequence 11, Appl
c 16	15.8	75.2		2503	3	US-08-461-145C-3	Sequence 3, Appli
c 17	15.8	75.2		2503	3	US-08-461-145C-11	Sequence 11, Appl
c 18	15.8	75.2		2503	4	US-08-628-829-5	Sequence 5, Appli
c 19	15.8	75.2		2503	4	US-08-628-829-7	Sequence 7, Appli
c 20	15.8	75.2		4729	6	5171684-1	Patent No. 5171684
c 21	15.8	75.2		6911	1	US-08-311-174-4	Sequence 4, Appli
c 22	15.8	75.2		7577	4	US-08-961-527-46	Sequence 46, Appl
c 23	15.2	72.4		399	4	US-09-641-638-45	Sequence 45, Appl
c 24	15.2	72.4		519	4	US-09-134-001C-1235	Sequence 1235, Ap
c 25	15.2	72.4		909	4	US-09-134-001C-334	Sequence 334, App
c 26	15.2	72.4		1084	2	US-08-184-009-110	Sequence 110, App
c 27	15.2	72.4		1084	2	US-08-458-356-110	Sequence 110, App
28	15.2	72.4		1084	4	US-08-460-736-110	Sequence 110, App
29	15.2	72.4		1094	2	US-08-184-009-109	Sequence 109, App
30	15.2	72.4		1094	2	US-08-458-356-109	Sequence 109, App
31	15.2	72.4		1094	4	US-08-460-736-109	Sequence 109, App
32	15.2	72.4		1691	2	US-08-993-118-8	Sequence 8, Appli
33	15.2	72.4		1691	2	US-08-845-528C-8	Sequence 8, Appli
34	15.2	72.4		1866	1	US-08-328-961-1	Sequence 1, Appli
35	15.2	72.4		1866	1	US-08-462-397-1	Sequence 1, Appli
36	15.2	72.4		2419	1	US-07-807-043B-7	Sequence 7, Appli
37	15.2	72.4		2419	1	US-08-299-849B-7	Sequence 7, Appli
38	15.2	72.4		2419	3	US-08-142-368A-7	Sequence 7, Appli
39	15.2	72.4		2419	4	US-08-967-727-7	Sequence 7, Appli
40	15.2	72.4		2419	4	US-08-037-230D-7	Sequence 7, Appli
41	15.2	72.4		2420	4	US-08-465-167A-23	Sequence 23, Appl
42	15.2	72.4		2420	4	US-09-056-105-4	Sequence 4, Appli
43	15.2	72.4		2420	4	US-08-627-820-23	Sequence 23, Appl
44	15.2	72.4		2511	2	US-08-363-560-1	Sequence 1, Appli
45	15.2	72.4		2848	2	US-08-805-918-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-961-527-48/c
; Sequence 48, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,527
;   FILING DATE:
;   CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 25002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; US-08-961-527-48

Query Match   78.1%; Score 16.4; DB 4; Length 25002;
Best Local Similarity 94.4%; Pred. No. 89;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   4 ATGGTGATGATTGGAATG 21
          ||| ||||||||||||||
Db   1372 ATGATGATGATTGGAATG 1355

RESULT 2
US-08-966-316-3
; Sequence 3, Application US/08966316
; Patent No. 5932445
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au-Young, Janice
; APPLICANT: Reddy, Roopa
; APPLICANT: Murry, Lynn E.
; APPLICANT: Mathur, Preete
; TITLE OF INVENTION: SIGNAL PEPTIDE - CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/966,316
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0424 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1141 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT15
; CLONE: 1682433
US-08-966-316-3

Query Match   77.1%; Score 16.2; DB 2; Length 1141;
Best Local Similarity   85.7%; Pred. No. 74;
Matches   18; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
         ||||||| ||||||| ||||||
Db   158 CTCATGATGATGACTGGATTG 178

RESULT 3
US-08-271-829-1
; Sequence 1, Application US/08271829
; Patent No. 5583021
; GENERAL INFORMATION:
; APPLICANT: William G. Dougherty and John A.
; APPLICANT: Lindbo
; TITLE OF INVENTION: Production of Virus
; TITLE OF INVENTION: Resistant Plants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Polley
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street, Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/271,829
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/838,509
; FILING DATE: February 19, 1992
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard J. Polley, Esq.
; REGISTRATION NUMBER: 28,107
; REFERENCE/DOCKET NUMBER: 245-40288
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9495
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to genomic RNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: Tobacco Etch Virus (TEV)
; STRAIN: Highly Aphid Transmitted
; STRAIN:
; IMMEDIATE SOURCE: TEV propagated in
; IMMEDIATE SOURCE: N.tabacum Burley 49
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: Coat protein gene
; LOCATION: Genomic nucleotides 8518
; LOCATION: -9306
; IDENTIFICATION METHOD: --
; OTHER INFORMATION: SEQ. ID No. 5583021 1 is
; OTHER INFORMATION: the cDNA corresponding to the Tobacco
; OTHER INFORMATION: Etch Virus Genome.
; PUBLICATION INFORMATION:
; AUTHORS: Allison et al.
; TITLE: The nucleotide sequence of the
; TITLE: Genomic RNA. Evidence for the   coding region o
; TITLE: Synthesis of a Single Polyprotein
; JOURNAL: Virology
; VOLUME: 154
; ISSUE: --
; PAGES: 9-20
US-08-271-829-1

Query Match   77.1%; Score 16.2; DB 1; Length 9495;
Best Local Similarity   85.7%; Pred. No. 97;
Matches   18; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
         ||||||| |||||||| ||||
Db 7942 CTCATGGTCATTATTGCAATG 7962

RESULT 4
PCT-US93-01544-1

; Sequence 1, Application PC/TUS9301544
; GENERAL INFORMATION:
; APPLICANT: William G. Dougherty and John A. Lindbo
; TITLE OF INVENTION: Production of Plants Showing
; TITLE OF INVENTION: Immunity to Viral Infection via Introduction of Genes
; TITLE OF INVENTION: Encoding Untranslatable Plus Sense RNA Molecules
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Polley
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street, Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01544
; FILING DATE: 19930218
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/838,509
; FILING DATE: 19 FEB 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard J. Polley, Esq.
; REGISTRATION NUMBER: 28,107
; REFERENCE/DOCKET NUMBER: 245-35829/RJP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9495
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to genomic RNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: Tobacco Etch Virus (TEV)
; STRAIN: Highly Aphid Transmitted (HAT)
; IMMEDIATE SOURCE: TEV propagated in N. tabacum
; IMMEDIATE SOURCE: Burley 49
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: Coat protein gene
; LOCATION: Genomic nucleotides 8518-9306
; IDENTIFICATION METHOD: --
; OTHER INFORMATION: SEQ. ID No. 1 is the cDNA
; OTHER INFORMATION: corresponding to the Tobacco Etch Virus Genome.
; PUBLICATION INFORMATION:
; AUTHORS: Allison et al.
; TITLE: The nucleotide sequence of the coding
; TITLE: region of Tobacco Etch Virus Genomic RNA:
; TITLE: Evidence for the Synthesis of a Single
; TITLE: Polyprotein
; JOURNAL: Virology
; VOLUME: 154
; ISSUE: --
; PAGES: 9-20
PCT-US93-01544-1

Query Match   77.1%; Score 16.2; DB 5; Length 9495;
Best Local Similarity 85.7%; Pred. No. 97; .
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 CTCATGGTTGATGATTGGAATG 21
            ||||| ||||| |||| ||||
Db   7942 CTCATGGTCATTATTGCAATG 7962

RESULT 5
US-08-961-527-99/c
; Sequence 99, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19718 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-99

Query Match   77.1%; Score 16.2; DB 4; Length 19718;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 CTCATGGTTGATGATTGGAATG 21
            || ||||||||| ||||||
Db   6079 CTTATGGTGATGAGTGGAAAG 6059

RESULT 6
US-08-858-207A-180/c
; Sequence 180, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328e1 Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

```
       COUNTRY: USA
       ZIP: 19406-0939
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ for Windows Version 2.0
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/858,207A
       FILING DATE: 09-MAY-1997
       CLASSIFICATION: 435
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER: 60/017670
       FILING DATE: 14-MAY-1996
     ATTORNEY/AGENT INFORMATION:
       NAME: Gimmi, Edward R
       REGISTRATION NUMBER: 38,891
       REFERENCE/DOCKET NUMBER: P50475
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: 610-270-4478
       TELEFAX: 610-270-5090
       TELEX:
   INFORMATION FOR SEQ ID NO: 180:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 836 base pairs
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
US-08-858-207A-180

Query Match        75.2%;  Score 15.8; DB 4;  Length 836;
Best Local Similarity  89.5%;  Pred. No. 1.1e+02;
Matches  17;  Conservative  0; Mismatches  2; Indels  0; Gaps  0

Qy     2 TCATGGTGATGATTGGAAT 20
          |||||| |||||||||||
Db   708 TCATGATGATGGTTGGAAT 690


RESULT 7
US-08-923-772-1/c
; Sequence 1, Application US/08923772
; Patent No. 5972651
; GENERAL INFORMATION:
;   APPLICANT: Black, Michael T.
;   TITLE OF INVENTION: NOVEL ffh
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Dechert Price & Rhoads
;     STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;     CITY: Philadelphia
;     STATE: PA
;     COUNTRY: US
;     ZIP: 19103
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/923,772
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Dickinson, Todd Q
;     REGISTRATION NUMBER: 28,354
;     REFERENCE/DOCKET NUMBER: GM10080
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 215-994-2252
;     TELEFAX: 215-994-2222
;     TELEX:
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1569 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
US-08-923-772-1

Query Match        75.2%;  Score 15.8; DB 2;  Length 1569;
Best Local Similarity  89.5%;  Pred. No. 1.2e+02;
Matches  17;  Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy     2 TCATGGTGATGATTGGAAT 20
          |||||| |||||||||||
Db   316 TCATGATGATGGTTGGAAT 298


RESULT 8
US-09-385-287-1/c
; Sequence 1, Application US/09385287
; Patent No. 6350857
; GENERAL INFORMATION:
;   APPLICANT: Black, Michael T.
;   TITLE OF INVENTION: NOVEL ffh
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Dechert Price & Rhoads
;     STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;     CITY: Philadelphia
;     STATE: PA
;     COUNTRY: US
;     ZIP: 19103
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/385,287
;     FILING DATE: 30-Aug-1999
;     CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/923,772
;     FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Dickinson, Todd Q
;     REGISTRATION NUMBER: 28,354
;     REFERENCE/DOCKET NUMBER: GM10080
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 215-994-2252
;     TELEFAX: 215-994-2222
;     TELEX: <Unknown>
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-385-287-1

   INFORMATION FOR SEQ ID NO: 1:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 1569 base pairs
       TYPE: nucleic acid
       STRANDEDNESS: double
       TOPOLOGY: linear
       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-385-287-1

Query Match        75.2%;  Score 15.8; DB 4;  Length 1569;
Best Local Similarity  89.5%;  Pred. No. 1.2e+02;
Matches  17;  Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy     2 TCATGGTGATGATTGGAAT 20
          |||||| |||||||||||
Db   316 TCATGATGATGGTTGGAAT 298


RESULT 9
US-09-423-890-3/c
```

; Sequence 3, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2013
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(1980)
US-09-423-890-3

Query Match   75.2%;   Score 15.8; DB 4;   Length 2013;
Best Local Similarity 89.5%;   Pred. No. 1.2e+02;
Matches   17;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAA 19
            ||| ||||||||| ||||
Db   951 CTCCTGGTGATGATAGGAA 933

RESULT 10
US-09-423-890-9/c
; Sequence 9, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2465
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (427)..(2283)
US-09-423-890-9

Query Match   75.2%;   Score 15.8; DB 4;   Length 2465;
Best Local Similarity 89.5%;   Pred. No. 1.2e+02;
Matches   17;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAA 19
            ||| ||||||||| ||||
Db   1254 CTCCTGGTGATGATAGGAA 1236

RESULT 11
US-08-472-934-3/c
; Sequence 3, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:

; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 466..2325
US-08-472-934-3

Query Match   75.2%;   Score 15.8; DB 1;   Length 2503;
Best Local Similarity 89.5%;   Pred. No. 1.3e+02;
Matches   17;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAA 19
            ||| ||||||||| ||||
Db   1293 CTCCTGGTGATGATAGGAA 1275

RESULT 12
US-08-472-934-11/c
; Sequence 11, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,934
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,421
FILING DATE: 15-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 5,405,941
FILING DATE: 15-Apr-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,460
FILING DATE: 14-Oct-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr. Esq.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: CPI-004DVCP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2503 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 466..2325
US-08-472-934-11

Query Match   75.2%;   Score 15.8; DB 1; Length 2503;
Best Local Similarity 89.5%;   Pred. No. 1.3e+02;
Matches   17; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAA 19
           ||| ||||||||| |||||
Db   1293 CTCCTGGTGATGATAGGAA 1275

RESULT 13
US-08-323-460A-3/c
; Sequence 3, Application US/08323460A
; Patent No. 5854043
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS & MCINTOSH
STREET: 1700 LINCOLN STREET, SUITE 3500
CITY: DENVER
STATE: CO
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,460A
FILING DATE: 14-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,254
FILING DATE: 14-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOVARIK, JOSEPH E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2503 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 466..2325
US-08-323-460A-3

Query Match   75.2%;   Score 15.8; DB 2; Length 2503;
Best Local Similarity 89.5%;   Pred. No. 1.3e+02;
Matches   17; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAA 19
           ||| ||||||||| |||||
Db   1293 CTCCTGGTGATGATAGGAA 1275

RESULT 14
US-08-461-146C-3/c
; Sequence 3, Application US/08461146C
; Patent No. 5981265
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,146C
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/354,516

```
;     FILING DATE:  14-OCT-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/US94/11690
;     FILING DATE:  14-OCT-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/US94/04178
;     FILING DATE:  15-APR-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  KARA, Catherine J.
;     REGISTRATION NUMBER:  P41,106
;     REFERENCE/DOCKET NUMBER:  CPI-004CN3
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (617) 227-7400
;     TELEFAX:  (617) 227-5941
; INFORMATION FOR SEQ ID NO:  11:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  2503 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  466..2325
US-08-461-146C-11

Query Match                  75.2%;   Score 15.8;  DB 2;   Length 2503;
Best Local Similarity   89.5%;   Pred. No. 1.3e+02;
Matches    17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 CTCATGGTGATGATTGGAA 19
            |||  ||||||||||| ||||
Db     1293 CTCCTGGTGATGATAGGAA 1275

Search completed: June  9, 2003, 11:07:52
Job time : 43.5 secs
```

```
;     FILING DATE:  21-FEB-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/049,254
;     FILING DATE:  15-APR-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/323,460
;     FILING DATE:  14-OCT-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/US94/11690
;     FILING DATE:  14-OCT-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/US94/04178
;     FILING DATE:  15-APR-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  KARA, Catherine J.
;     REGISTRATION NUMBER:  P41,106
;     REFERENCE/DOCKET NUMBER:  CPI-004CN3
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (617) 227-7400
;     TELEFAX:  (617) 227-5941
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  2503 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  466..2325
US-08-461-146C-3

Query Match                  75.2%;   Score 15.8;  DB 2;   Length 2503;
Best Local Similarity   89.5%;   Pred. No. 1.3e+02;
Matches    17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 CTCATGGTGATGATTGGAA 19
            |||  ||||||||||| ||||
Db     1293 CTCCTGGTGATGATAGGAA 1275

RESULT 15
US-08-461-146C-11/c
; Sequence 11, Application US/08461146C
; Patent No. 5581265
; GENERAL INFORMATION:
;   APPLICANT:  JOHNSON, GARY L.
;   TITLE OF INVENTION:  METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
;   NUMBER OF SEQUENCES:  14
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Lahive and Cockfield
;     STREET:  60 State Street
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  U.S.A.
;     ZIP:  02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/461,146C
;     FILING DATE:  05-JUN-1995
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/354,516
;     FILING DATE:  21-FEB-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/049,254
;     FILING DATE:  15-APR-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/323,460
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 09:12:21 ; Search time 59.5 Seconds
                                     (without alignments)
                                     493.954 Million cell updates/sec

Title:     US-09-550-163-1_COPY_230_250
Perfect score: 21
Sequence:  1 ctcatggtgatgatggaatg 21

Scoring table:  IDENTITY NUC
                Gapop 10.0, Gapext 1.0

Searched:  870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters:        1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published Applications NA:*
         1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
         2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
         3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
         4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
         5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
         7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
         8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
         9:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
        10:  /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
        11:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        12:  /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
        13:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
        14:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	100.0	312	10	US-09-864-761-20233	Sequence 20233, A
	2	21	100.0	372	10	US-09-864-761-33139	Sequence 33139, A
	3	21	100.0	450	10	US-09-864-761-3463	Sequence 3463, Ap
	4	21	100.0	471	10	US-09-864-761-16671	Sequence 16671, A
	5	21	100.0	732	9	US-10-000-151B-5	Sequence 5, Appli
	6	21	100.0	113604	9	US-10-227-195A-1	Sequence 1, Appli
	7	21	100.0	113604	9	US-10-227-195A-2	Sequence 2, Appli
	8	16.8	80.0	583	10	US-09-864-761-20772	Sequence 20772, A
c	9	16.8	80.0	707	9	US-10-011-585A-78	Sequence 78, Appl
	10	16.8	80.0	1959	10	US-09-864-761-4012	Sequence 4012, Ap
c	11	16.4	78.1	1350	9	US-09-769-787-322	Sequence 322, App
	12	16.4	78.1	1448	9	US-09-774-639-63	Sequence 63, Appl
	13	16.4	78.1	1448	9	US-09-969-730-52	Sequence 52, Appl
	14	16.2	77.1	227	10	US-09-294-093B-5363	Sequence 5363, Ap
c	15	16.2	77.1	481	9	US-09-968-433-26	Sequence 26, Appl
c	16	16.2	77.1	574	10	US-09-864-761-12791	Sequence 12791, A
	17	16.2	77.1	1141	9	US-09-968-433-3	Sequence 3, Appli
	18	16.2	77.1	1653	10	US-09-529-063-80	Sequence 80, Appl
	19	16.2	77.1	2000	10	US-09-529-063-81	Sequence 81, Appl
	20	16.2	77.1	96649	9	US-09-956-712-10	Sequence 10, Appl
c	21	16.2	77.1	536165	9	US-09-939-964-1	Sequence 1, Appli
c	22	16	76.2	640681	10	US-09-790-988-1	Sequence 1, Appli
	23	15.8	75.2	257	10	US-09-878-574-7121	Sequence 7121, Ap
	24	15.8	75.2	271	10	US-09-878-574-12129	Sequence 12129, A
	25	15.8	75.2	382	10	US-09-878-574-3431	Sequence 3431, Ap
	26	15.8	75.2	774	10	US-09-770-445-898	Sequence 898, App
	27	15.8	75.2	795	9	US-09-822-846-314	Sequence 314, App
	28	15.8	75.2	1299	9	US-09-738-626-2905	Sequence 2905, Ap
	29	15.8	75.2	1395	10	US-09-815-242-9427	Sequence 9427, Ap
	30	15.8	75.2	1524	9	US-09-938-842A-300	Sequence 300, App
	31	15.8	75.2	1569	10	US-09-814-041A-1	Sequence 1, Appli
	32	15.8	75.2	1572	10	US-09-815-242-9156	Sequence 9156, Ap
	33	15.8	75.2	1575	10	US-09-981-353-75	Sequence 75, Appli
	34	15.8	75.2	2013	12	US-10-000-864-3	Sequence 3, Appli
	35	15.8	75.2	2465	12	US-10-000-864-9	Sequence 9, Appli
	36	15.8	75.2	6504	9	US-09-801-368-55	Sequence 55, Appl
	37	15.8	75.2	3309400	9	US-09-738-626-1	Sequence 1, Appli
	38	15.4	73.3	2000	9	US-09-938-842A-4825	Sequence 4825, Ap
	39	15.2	72.4	109	10	US-09-864-761-17878	Sequence 17878, A
c	40	15.2	72.4	167	10	US-09-960-352-6504	Sequence 6504, Ap
	41	15.2	72.4	213	10	US-09-864-761-29571	Sequence 29571, A
	42	15.2	72.4	260	10	US-09-878-574-6372	Sequence 6372, Ap
	43	15.2	72.4	335	9	US-09-782-974C-53	Sequence 53, Appl
c	44	15.2	72.4	340	9	US-09-984-130-106	Sequence 106, App
c	45	15.2	72.4	346	10	US-09-878-574-3002	Sequence 3002, Ap

ALIGNMENTS

```
RESULT 1
US-09-864-761-20233
; Sequence 20233, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT: Rank, David R.
;  APPLICANT: Hanzel, David K.
;  APPLICANT: Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR APPLICATION FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR APPLICATION FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR APPLICATION FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR APPLICATION FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR APPLICATION FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR APPLICATION FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR APPLICATION FILING DATE: 2001-01-30
```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20233
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000052.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.88
; OTHER INFORMATION: EST_HUMAN HIT: AI246239.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y6J6, EVALUE 3.00e-55
; OTHER INFORMATION: NT HIT: AF302095.1, EVALUE 0.00e+00
US-09-864-761-20233

Query Match   100.0%;   Score 21; DB 10;   Length 312;
Best Local Similarity   100.0%;   Pred. No. 2;
Matches   21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   104 CTCATGGTGATGATTGGAATG 124

RESULT 2
US-09-864-761-33139
; Sequence 33139, Application US/09864761
; Patent No. US2002004B763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33139
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000120.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
; OTHER INFORMATION: EST_HUMAN HIT: AI654552.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y6J6, EVALUE 8.00e-67
; OTHER INFORMATION: NT HIT: gi11526220, EVALUE 0.00e+00
US-09-864-761-33139

Query Match   100.0%;   Score 21; DB 10;   Length 372;
Best Local Similarity   100.0%;   Pred. No. 2.1;
Matches   21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   157 CTCATGGTGATGATTGGAATG 177

RESULT 3
US-09-864-761-3463
; Sequence 3463, Application US/09864761
; Patent No. US2002004B763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3463
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000052.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.88
US-09-864-761-3463

Query Match   100.0%; Score 21; DB 10; Length 450;
Best Local Similarity   100.0%; Pred. No. 2.1;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   312 CTCATGGTGATGATTGGAATG 332

RESULT 4
US-09-864-761-16671
; Sequence 16671, Application US/09864761
; Patent No. US2002004876A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16671
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000120.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
US-09-864-761-16671

Query Match   100.0%; Score 21; DB 10; Length 471;
Best Local Similarity   100.0%; Pred. No. 2.1;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   395 CTCATGGTGATGATTGGAATG 415

RESULT 5
US-10-000-151B-5
; Sequence 5, Application US/10000151B
; Publication No. US20030011136A1
; GENERAL INFORMATION:
; APPLICANT: Balser, Jeffrey R.
; APPLICANT: George, Alfred L.
; FILE REFERENCE: Vanderbilt Ref No. US2003001136A1 VU0120; Attorney Docket No. US200
; TITLE OF INVENTION: HUMAN KCR1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; CURRENT APPLICATION NUMBER: US/10/000,151B
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-000-151B-5

Query Match   100.0%; Score 21; DB 9; Length 732;
Best Local Similarity   100.0%; Pred. No. 2.3;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              |||||||||||||||||||||
Db   230 CTCATGGTGATGATTGGAATG 250

RESULT 6
US-10-227-195A-1
; Sequence 1, Application US/10227195A
; Publication No. US20030077633A1
; GENERAL INFORMATION:
;   APPLICANT: Cox, David
;   APPLICANT: Arnold, Deana
;   TITLE OF INVENTION: Haplotype structure of chromosome 21
;   FILE REFERENCE: 1030U1
;   CURRENT APPLICATION NUMBER: US/10/227,195A
;   CURRENT FILING DATE: 2002-11-18
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 113604
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 7175, 7204, 36973, 66372, 76921, 81512, 88727
;   OTHER INFORMATION: n = G or C
US-10-227-195A-1

Query Match   100.0%; Score 21; DB 9; Length 113604;
Best Local Similarity   100.0%; Pred. No. 4.6;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              ||||||||||||||||||||
Db   17632 CTCATGGTGATGATTGGAATG 17652

RESULT 7
US-10-227-195A-2
; Sequence 2, Application US/10227195A
; Publication No. US20030077633A1
; GENERAL INFORMATION:
;   APPLICANT: Cox, David
;   APPLICANT: Arnold, Deana
;   TITLE OF INVENTION: Haplotype structure of chromosome 21
;   FILE REFERENCE: 1030U1
;   CURRENT APPLICATION NUMBER: US/10/227,195A
;   CURRENT FILING DATE: 2002-11-18
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 113604
;   TYPE: DNA
;   ORGANISM: Human
US-10-227-195A-2

Query Match   100.0%; Score 21; DB 9; Length 113604;
Best Local Similarity   100.0%; Pred. No. 4.6;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
              ||||||||||||||||||||
Db   17632 CTCATGGTGATGATTGGAATG 17652

RESULT 8
US-09-864-761-20772
; Sequence 20772, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
;   APPLICANT: Penn, Sharron G.
;   APPLICANT: Rank, David R.
;   APPLICANT: Hanzel, David K.
;   APPLICANT: Chen, Wensheng
;   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;   TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;   FILE REFERENCE: Aeomica-X-1
;   CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20772
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022334.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
;   OTHER INFORMATION: EST_HUMAN HIT: AW992395.1, EVALUE 9.50e-02
US-09-864-761-20772

Query Match   80.0%; Score 16.8; DB 10; Length 583;
Best Local Similarity   90.0%; Pred. No. 1.9e+02;
Matches   18; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
              |||||||||||| ||| ||||
Db   31 TGATGGTGATGATGGGAATG 50

RESULT 9
US-10-011-585A-78
; Sequence 78, Application US/10011585A
; Publication No. US2003039986A1
; GENERAL INFORMATION:
;   APPLICANT: Sun, Yongming

; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0261
; CURRENT APPLICATION NUMBER: US/10/011,585A
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/245,740
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 245
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (307)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (359)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (407)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (421)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (434)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (436)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (477)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (482)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (484)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (521)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (555)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (561)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (583)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (606)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (618)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (679)
; OTHER INFORMATION: a, c, g or t
; NAME/KEY: unsure
; LOCATION: (700)
; OTHER INFORMATION: a, c, g or t
US-10-011-585A-78

Query Match   80.0%;   Score 16.8;   DB 9;   Length 707;
Best Local Similarity   90.0%;   Pred. No. 1.9e+02;
Matches   18;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 CTCATGGTGATGATTGGAAT 20
            |||||||||| ||| |||||
Db   630 CTCATGGTGAAGATAGGAAT 649

RESULT 10
US-09-864-761-4012
; Sequence 4012, Application US/09864761
; Patent No. US2002004B763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERNCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4012
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022334.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
US-09-864-761-4012

Query Match   80.0%;   Score 16.8; DB 10;   Length 1959;
Best Local Similarity 90.0%;   Pred. No. 2.2e+02;
Matches 18; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
          |||||||||||||| |||||
Db   536 TGATGGTGATGATGGGAATG 555

RESULT 11
US-09-769-787-322/c
; Sequence 322, Application US/09769787
; Publication No. US2003091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 322
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-322

Query Match   78.1%;   Score 16.4; DB 9;   Length 1350;
Best Local Similarity 94.4%;   Pred. No. 3.2e+02;
Matches 17; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   4 ATGGTGATGATTGGAATG 21
          ||| |||||||||||||||
Db 1303 ATGATGATGATTGGAATG 1286

RESULT 12
US-09-774-639-63
; Sequence 63, Application US/09774639
; Publication No. US2003003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: PZ013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1448
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-774-639-63

Query Match   78.1%;   Score 16.4; DB 9;   Length 1448;
Best Local Similarity 94.4%;   Pred. No. 3.3e+02;
Matches 17; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   2 TCATGGTGATGATTGGAA 19
          |||||||||| |||||||
Db 1094 TCATGGTGATAATTGGAA 1111

RESULT 13
US-09-969-730-52
; Sequence 52, Application US/09969730
; Publication No. US2003005443A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: PZ013P2
; CURRENT APPLICATION NUMBER: US/09/969,730
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,367
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,365
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,731
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,557
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,563
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,970
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,986
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,311
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,808
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,803
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,809
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,806
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,310
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,798
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,309
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,312
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,807
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,386
; PRIOR FILING DATE: 1997-08-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1448
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-969-730-52

Query Match   78.1%; Score 16.4; DB 9; Length 1448;
Best Local Similarity   94.4%; Pred. No. 3.3e+02;
Matches   17; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   2 TCATGGTGATGATTGGAA 19
               ||||||||| |||||||||
Db   1094 TCATGGTGATAATTGGAA 1111

RESULT 14
US-09-294-093B-5363/c
; Sequence 5363, Application US/09294093B
; Patent No. US2001005135A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5363
;   LENGTH: 227
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US2001005135A1 70035460H1
;   NAME/KEY: unsure
;   LOCATION: 195, 206
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5363

Query Match   77.1%; Score 16.2; DB 10; Length 227;
Best Local Similarity   85.7%; Pred. No. 3.1e+02;
Matches   18; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
               ||||||||| ||| ||||||
Db   193 CTGATGGTGAGGATGGGAATG 173

RESULT 15
US-09-968-433-26
; Sequence 26, Application US/09968433
; Publication No.US20030073162A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au-Young, Janice
; APPLICANT: Reddy, Roopa
; APPLICANT: Murry, Lynn E.
; APPLICANT: Mathur, Preete
; TITLE OF INVENTION: SIGNAL PEPTIDE-CONTAINING PROTEINS
; FILE REFERENCE: PC-0051 CIP
; CURRENT APPLICATION NUMBER: US/09/968,433
; CURRENT FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PERL Program
; SEQ ID NO 26
;   LENGTH: 481
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030073162A1 2444714F6
US-09-968-433-26

Query Match   77.1%; Score 16.2; DB 9; Length 481;
Best Local Similarity   85.7%; Pred. No. 3.5e+02;
Matches   18; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 CTCATGGTGATGATTGGAATG 21
               ||||||||| |||| ||||||
Db   158 CTCATGATGAGACTGGATTG 178

Search completed: June 9, 2003, 12:10:20
Job time : 71.5 secs

```
                        GenCore version 5.1.6
               Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 08:36:50 ; Search time 883 Seconds
                                        (without alignments)
                                   385.170 Million cell updates/sec

Title:        US-09-550-163-1_COPY_230_250
Perfect score: 21
Sequence:     1 ctcatggtgatgatggaatg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:     16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:     32308112

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
          1:  em_estba:*
          2:  em_esthum:*
          3:  em_estin:*
          4:  em_estmu:*
          5:  em_estov:*
          6:  em_estpl:*
          7:  em_estro:*
          8:  em_htc:*
          9:  gb_est1:*
          10: gb_est2:*
          11: gb_htc:*
          12: gb_est3:*
          13: gb_est4:*
          14: gb_est5:*
          15: em_estfun:*
          16: em_estom:*
          17: gb_gss:*
          18: em_gss_hum:*
          19: em_gss_inv:*
          20: em_gss_pln:*
          21: em_gss_vrt:*
          22: em_gss_fun:*
          23: em_gss_mam:*
          24: em_gss_mus:*
          25: em_gss_other:*
          26: em_gss_pro:*
          27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	21	100.0	121	9	AA935321	AA935321 oo71g09.s
2	21	100.0	311	10	AW869303	AW869303 MR3-SN006
3	21	100.0	372	9	AI246239	AI246239 qi29g04.x
4	21	100.0	391	9	AI339609	AI339609 qq42a07.x
5	21	100.0	410	9	AI962650	AI962650 wq42e03.x
6	21	100.0	429	9	AI654552	AI654552 wb48b12.x
c 7	21	100.0	803	12	BG208163	BG208163 RST27654
c 8	18.4	87.6	478	13	BM266405	BM266405 MEST380-D
c 9	18.4	87.6	532	13	BM337949	BM337949 MEST219-F
c 10	18.4	87.6	542	10	AW066747	AW066747 683008B08
c 11	18.4	87.6	579	13	AW091474	AW091474 683028H09
c 12	18.4	87.6	595	13	BM351227	BM351227 MEST275-A
c 13	18.4	87.6	610	10	AW076451	AW076451 683018E04
c 14	18.4	87.6	623	13	AV702181	AV702181 AV702181
c 15	18.4	87.6	634	13	BM351807	BM351807 MEST346-G
c 16	18.4	87.6	638	13	BM336865	BM336865 MEST199-H
c 17	18.4	87.6	661	13	BM336951	BM336951 MEST201-C
c 18	18.4	87.6	714	13	BM339352	BM339352 MEST240-C
c 19	17.8	84.8	351	13	BG938225	BG938225 1Ab14A12
c 20	17.8	84.8	439	17	BH302773	BH302773 CH230-97O
c 21	17.8	84.8	470	14	D85797	D85797 D85797 Rat
c 22	17.8	84.8	488	13	BM512971	BM512971 kx82h01.y
c 23	17.8	84.8	581	17	B94648	B94648 CIT-HSP-216
c 24	17.8	84.8	589	13	B1787454	B1787454 sai45h11
c 25	17.8	84.8	903	17	A2679050	A2679050 ENTMI13TF
c 26	17.8	84.8	910	17	A2676540	A2676540 ENTUN10TF
c 27	17.8	84.8	1003	12	BG261965	BG261965 Mus muscu
c 28	17.8	84.8	1691	11	AK008619	AK008619 Mus muscu
c 29	17.4	82.9	513	11	A2019690	A2019690 RPCI-23-2
c 30	17.4	82.9	578	14	BQ473783	BQ473783 sap1h04.
c 31	17.4	82.9	601	10	BE248115	BE248115 NF003G12D
c 32	17.4	82.9	663	12	BG450315	BG450315 NF011D12D
c 33	17.4	82.9	668	12	BF632575	BF632575 NF012A01D
c 34	17	81.0	489	14	H95139	H95139 yw59c08.s1
c 35	17	81.0	746	12	BG221966	BG221966 RST41783
c 36	17	81.0	905	12	BG114301	BG114301 602286063
c 37	16.8	80.0	222	14	BQ526572	BQ526572 NISC_no17
c 38	16.8	80.0	231	17	BH872665	BH872665 hp07e09.b
c 39	16.8	80.0	238	14	T18753	T18753 2C02A04-T7
c 40	16.8	80.0	277	14	BQ333632	BQ333632 MR4-PT005
c 41	16.8	80.0	277	17	AQ482802	AQ482802 RPCI-11-2
c 42	16.8	80.0	307	17	BJ269854	BJ269854
c 43	16.8	80.0	353	9	AA761065	AA761065 ny13d09.s
c 44	16.8	80.0	368	12	BF754132	BF754132 CM4-CT051
c 45	16.8	80.0	371	13	BG931453	BG931453 f33-88 f3

ALIGNMENTS

```
RESULT 1
AA935321
AA935321
LOCUS         AA935321    121 bp    mRNA    linear   EST 07-JUL-1998
DEFINITION    oo71g09.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1571680 3'
              similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
              CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION     AA935321
VERSION       AA935321.1    GI:3092478
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 121)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
```

FEATURES
    source
Trace considered overall poor quality
Insert Length: 763 Std Error: 0.00
Seq primer: -40m13 fwd ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..121
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:157180"
/clone_lib="NCI_CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo. "

BASE COUNT   33 a   31 c   28 g   29 t
ORIGIN

Query Match   100.0%; Score 21; DB 9; Length 121;
Best Local Similarity   100.0%; Pred. No. 19;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy 1 CTCATGGTGATGATTGGAATG 21
        |||||||||||||||||||||
Db 5 CTCATGGTGATGATTGGAATG 25

RESULT 2
AW869303
LOCUS   AW869303   311 bp   mRNA   linear   EST 22-MAY-2000
DEFINITION MR3-SN0067-240400-006-f11 SN0067 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW869303
VERSION   AW869303.1 GI:8003356
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 311)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT   Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-SN0067-240
            400-006-f11&t3=2000-04-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 74.
FEATURES   Location/Qualifiers
    source   1..311
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"

/clone_lib="SN0067"
/dev_stage="Adult"
/note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT   60 a   80 c   76 g   95 t
ORIGIN

Query Match   100.0%; Score 21; DB 10; Length 311;
Best Local Similarity   100.0%; Pred. No. 28;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy 1 CTCATGGTGATGATTGGAATG 21
        |||||||||||||||||||||
Db 25 CTCATGGTGATGATTGGAATG 45

RESULT 3
AI246239
LOCUS   AI246239   372 bp   mRNA   linear   EST 28-JAN-1999
DEFINITION qi29g04.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1857942
            3' similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI246239
VERSION   AI246239.1 GI:3841636
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 372)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 921   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 365.
FEATURES   Location/Qualifiers
    source   1..372
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1857942"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."

BASE COUNT   115 a   89 c   86 g   82 t
ORIGIN

Query Match   100.0%; Score 21; DB 9; Length 372;
Best Local Similarity   100.0%; Pred. No. 31;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy       1 CTCATGGTGATGATTGGAATG 21
           ||||||||||||||||||||||
Db     113 CTCATGGTGATGATTGGAATG 133

RESULT 4
AI339609
LOCUS       AI339609       391 bp    mRNA    linear    EST 29-DEC-1998
DEFINITION  qq42a07.x1 Soares NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1935156
            3' similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI339609
VERSION     AI339609.1  GI:4076536
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 391)
AUTHORS     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE       Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 380.
FEATURES             Location/Qualifiers
     source          1..391
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1935156"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT    119 a   93 c   93 g   86 t
ORIGIN

Query Match           100.0%;   Score 21;  DB 9;  Length 391;
Best Local Similarity 100.0%;   Pred. No. 31;
Matches      21;  Conservative 0; Mismatches  0; Indels  0; Gaps  0;

Qy       1 CTCATGGTGATGATTGGAATG 21
           ||||||||||||||||||||||
Db     122 CTCATGGTGATGATTGGAATG 142

RESULT 5
AI962650
LOCUS       AI962650       410 bp    mRNA    linear    EST 08-MAR-2000
DEFINITION  wq42e03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2473948 3'
            similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI962650
VERSION     AI962650.1  GI:5755363
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


REFERENCE
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS     1 (bases 1 to 410)
TITLE       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL     Tumor Gene Index
            Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 770   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2473948"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I, Site_2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP GC4 was prepared, and
                     ss circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT    120 a   93 c   93 g   104 t
ORIGIN

Query Match           100.0%;   Score 21;  DB 9;  Length 410;
Best Local Similarity 100.0%;   Pred. No. 32;
Matches      21;  Conservative 0; Mismatches  0; Indels  0; Gaps  0;

Qy       1 CTCATGGTGATGATTGGAATG 21
           ||||||||||||||||||||||
Db     138 CTCATGGTGATGATTGGAATG 158

RESULT 6
AI654552
LOCUS       AI654552       429 bp    mRNA    linear    EST 17-DEC-1999
DEFINITION  wb48b12.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2308895 3'
            similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI654552
VERSION     AI654552.1  GI:4738531
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 429)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
```

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 771   Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 411.

FEATURES   Location/Qualifiers
     source   1..429
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2308895"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I, Site_2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP_GC4 was prepared, and
                     as circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT   127 a   100 c   97 g   104 t   1 others
ORIGIN

Query Match   100.0%; Score 21; DB 9; Length 429;
Best Local Similarity   100.0%; Pred. No. 33;
Matches 21;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   132 CTCATGGTGATGATTGGAATG 152

RESULT 7
BG208163/c
LOCUS   BG208163   803 bp   mRNA   linear   EST 21-APR-2001
DEFINITION RST27654 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG208163
VERSION   BG208163.1 GI:13729850
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 803)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE   Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
   MEDLINE   21227151
  COMMENT   Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 550.

FEATURES   Location/Qualifiers
     source   1..803
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the

---

cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

BASE COUNT   222 a   172 c   171 g   238 t
ORIGIN

Query Match   100.0%; Score 21; DB 12; Length 803;
Best Local Similarity   100.0%; Pred. No. 42;
Matches 21;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 CTCATGGTGATGATTGGAATG 21
            |||||||||||||||||||||
Db   531 CTCATGGTGATGATTGGAATG 511

RESULT 8
BM266405/c
LOCUS   BM266405   478 bp   mRNA   linear   EST 02-JUL-2002
DEFINITION MEST380-D03.T3 ISUM5-RN Zea mays cDNA clone ME80-D03 3', mRNA
            sequence.
ACCESSION   BM266405
VERSION   BM266405.1 GI:17929445
KEYWORDS   EST.
SOURCE   Zea mays.
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 478)
  AUTHORS   Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
  TITLE   Expressed Sequence Tags from B73 Maize: various stages and tissues
            including seedlings treated with a variety of hormones
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Patrick S. Schnable
            Schnable Laboratory
            Iowa State University
            G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
            Tel: 515-294-0975
            Fax: 515-294-2299
            Email: schnable@iastate.edu
            Individual basecall and confidence value were assigned using the
            Phred software,
            (<http://depts.washington.edu/ventures/collabtr/direct/index.htm>#b
            rt). Overall sequence quality assessment and vector trimming were
            conducted using the Lucy software (<http://www.tigr.org/softlab/>).
            Lucy parameters were set to ensure an overall trimmed quality of
            97.5% or better without any vector fragments in the chosen
            high-quality region of each sequence. Low-quality bases between the
            poly-T and the high-quality region were replaced with N's to serve
            as spacers.
            PCR PRimers
            FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)
            BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)
            Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES   Location/Qualifiers
     source   1..478
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ME80-D03"
                     /clone_lib="ISUM5-RN"
                     /tissue_type="mixed"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
                     Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
                     Mixed mature tissures (17, 21, 38, 69, 77 DAG), Kernels
                     (3, 5, 10, 15, 20, 25, 30, DAP), Adventious roots (65 DAG
                     ), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0
                     cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
                     first ear, ear shank, etiolated seedlings, callus,
                     Cycloheximide-treated callus, Anaerobic treated seedlings
                     NAA (a-Naphthalene acetic acid)-treated seedlings,
                     Kinetin-treated seedlings, ACPC

(1-aminocyclopropane-1-carboxylix acid)-treated seedlings,
Brassinolide-treated seedlings, ABA (Abscisic acid
)-treated seedlings, GA (Gibberellic acid)-treated
seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA
molecules were generated as follows. First-strand cDNA was
prepared from oligo-dT selected mRNA by priming with a
NotI oligo-dT primer (5'
AACTGGAAGAATTCGGCGCCGCCAGGAATTTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA Poli-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT7T3PAC vector. The library
then went through one round of normalization to CoT value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

BASE COUNT   101 a   117 c   133 g   127 t
ORIGIN

Query Match   87.6%;   DB 13;   Length 478;
Best Local Similarity   95.0%;   Pred. No. 4.9e+02;
Matches   19;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
           |||||||||||||| |||||
Db   100 TCATGGTGATGATAGGAATG 81

RESULT 9
BM337949/c
LOCUS   BM337949   532 bp   mRNA   linear   EST 16-JAN-2002
DEFINITION   MEST219-F04.T3 ISUM5-RN Zea mays cDNA clone MEST219-F04 3', mRNA
                sequence.
ACCESSION   BM337949
VERSION   BM337949.1 GI:18168109
KEYWORDS   EST.
SOURCE   Zea mays.
ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 532)
AUTHORS   Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
TITLE   Expressed Sequence Tags from B73 Maize: various stages and tissues
                including seedlings treated with a variety of hormones
JOURNAL   Unpublished (2001)
COMMENT   Contact: Patrick S. Schnable
                Schnable Laboratory
                Iowa State University
                G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
                Tel: 515-294-0975
                Fax: 515-294-2299
                Email: schnable@iastate.edu
                Individual basecall and confidence value were assigned using the
                Phred software.
                (<http://deps.washington.edu/ventures/collabtr/direct/index.htm>#b
                rt). Overall sequence quality assessment and vector trimming were
                conducted using the Lucy software (<http://www.tigr.org/softlab/>).
                Lucy parameters were set to ensure an overall trimmed quality of
                97.5% or better without any vector fragments in the chosen
                high-quality region of each sequence. Low-quality bases between the
                poly-T and the high-quality region were replaced with N's to serve
                as spacers.
                PCR PRimers
                FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)
                BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)
                Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
FEATURES   Location/Qualifiers
     source   1..532
                    /organism="Zea mays"
                    /cultivar="B73"
                    /db_xref="taxon:4577"

/clone="MEST219-F04"
/clone_lib="ISUM5-RN"
/tissue_type="mixed"
/lab_host="DH10B"
/note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels
(3, 5, 10, 15, 20, 25, 30, DAP), Adventious roots (65 DAG
), Tassel (1-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0
cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
first ear, ear shank, etiolated seedlings, callus,
Cycloheximide-treated callus, Anaerobic treated seedlings
, NAA (a-Naphthalene acetic acid)-treated seedlings,
Kinetin-treated seedlings, ACPC
(1-aminocyclopropane-1-carboxylix acid)-treated seedlings,
Brassinolide-treated seedlings, ABA (Abscisic acid
)-treated seedlings, GA (Gibberellic acid)-treated
seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA
molecules were generated as follows. First-strand cDNA was
prepared from oligo-dT selected mRNA by priming with a
NotI oligo-dT primer (5'
AACTGGAAGAATTCGGCGCCGCCAGGAATTTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA Poli-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT7T3PAC vector. The library
then went through one round of normalization to CoT value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

BASE COUNT   120 a   130 c   145 g   137 t
ORIGIN

Query Match   87.6%;   DB 13;   Length 532;
Best Local Similarity   95.0%;   Pred. No. 5.1e+02;
Matches   19;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
           |||||||||||||| |||||
Db   123 TCATGGTGATGATAGGAATG 104

RESULT 10
AW066747/c
LOCUS   AW066747   542 bp   mRNA   linear   EST 12-OCT-1999
DEFINITION   683008B08.x1 683 - 14 day immature embryo from Hake lab (HS) Zea
                mays cDNA, mRNA sequence.
ACCESSION   AW066747
VERSION   AW066747.1 GI:6021819
KEYWORDS   EST.
SOURCE   Zea mays.
ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 542)
AUTHORS   Walbot.V.
TITLE   Maize ESTs from various cDNA libraries sequenced at Stanford
                University
JOURNAL   Unpublished (1999)
COMMENT   Contact: Walbot V
                Department of Biological Sciences
                Stanford University
                855 California Ave, Palo Alto, CA 94304, USA
                Tel: 650 723 2227
                Fax: 650 725 8221
                Email: walbot@stanford.edu
                Plate: 683008 row: B column: 08.
FEATURES   Location/Qualifiers
     source   1..542
                    /organism="zea mays"
                    /cultivar="B73"

RESULT 12
BM351227/c
LOCUS   BM351227   MEST275-A06.T3 ISUMS-RN Zea mays cDNA clone MEST275-A06 3', mRNA
DEFINITION   MEST275-A06.T3 ISUMS-RN Zea mays cDNA clone MEST275-A06 3', mRNA
               sequence.
ACCESSION   BM351227
VERSION   BM351227.1 GI:1817B183
KEYWORDS   EST.
SOURCE   Zea mays.
ORGANISM   Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 595)
AUTHORS   Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
TITLE   Expressed Sequence Tags from B73 Maize: various stages and tissues
               including seedlings treated with a variety of hormones
JOURNAL   Unpublished (2001)
COMMENT   Contact: Patrick S. Schnable
               Schnable Laboratory
               Iowa State University
               G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
               Tel: 515-294-0975
               Fax: 515-294-2299
               Email: schnable@iastate.edu
               Individual basecall and confidence value were assigned using the
               Phred software.
               (<http://depts.washington.edu/ventures/collabr/direct/index.htm>#b
               rt). Overall sequence quality assessment and vector trimming were
               conducted using the lucy software (<http://www.tigr.org/softlab/>).
               Lucy parameters were set to ensure an overall trimmed quality of
               97.5% or better without any vector fragments in the chosen
               high-quality region of each sequence. Low-quality bases between the
               poly-T and the high-quality region were replaced with N's to serve
               as spacers.
               PCR Primers
               FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)
               BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)
               Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
FEATURES   Location/Qualifiers
source   1..595
               /organism="Zea mays"
               /cultivar="B73"
               /db_xref="taxon:4577"
               /clone="MEST275-A06"
               /clone_lib="ISUMS-RN"
               /tissue_type="mixed"
               /lab_host="DH10B"
               /note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
               Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
               Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels
               (3, 5, 10, 15, 20, 25, 30, DAP), Adventious roots (65 DAG
               ), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0
               cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
               first ear, ear shank, etiolated seedlings, callus,
               Cycloheximide-treated callus, Anaerobic treated seedlings
               , NAA (a-Naphthalene acetic acid)-treated seedlings,
               Kinetin-treated seedlings, ACPC
               (1-aminocyclopropane-1-carboxylix acid)-treated seedlings,
               Brassinolide-treated seedlings, ABA (Abscisic acid
)-treated seedlings, GA (Gibberellic acid)-treated
               seedlings, JA (Jasmonic acid)-treated seedlings, ds-cDNA
               molecules were generated as follows. First-strand cDNA was
               prepared from oligo-dT selected mRNA by priming with a
               NotI oligo-dT primer (5'
               AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT). The
               resulting DNA:RNA hybrid was treated with RNase H and used
               as a template for DNA PolI-catalyzed second strand
               synthesis. After the addition of EcoRI adaptors, the
               ds-cDNAs were digested with NotI and size-selected. The
               resulting molecules were directionally cloned into the
               EcoRI and NotI sites of the pT7T3PAC vector. The library

               /db_xref="taxon:4577"
               /clone_lib="DH10B"
)"
               /tissue_type="embryo"
               /dev_stage="14 days after pollination"
               /lab_host="DH10B"
               /note="Organ: embryo; Vector: pBKCMV (Stratagene's Zap
               Express); Site_1: XhoI; Site_2: EcoRI; Directionally
               cloned, 14 day immature embryo library created with
               Stratagene's Zap Express cDNA protocol."
BASE COUNT   121 a   131 c   146 g   143 t   1 others
ORIGIN

Query Match   87.6%; Score 18.4; DB 10; Length 542;
Best Local Similarity 95.0%; Pred. No. 5.2e+02;
Matches   19; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
        |||||||||||||||| ||||
Db 132 TCATGGTGATGATAGGAATG 113

RESULT 11
AW091474/c
LOCUS   AW091474   579 bp   mRNA   linear   EST 18-OCT-1999
DEFINITION   683028H09.x1 683 - 14 day immature embryo from Hake lab (HS) Zea
               mays cDNA, mRNA sequence.
ACCESSION   AW091474
VERSION   AW091474.1 GI:6057069
KEYWORDS   EST.
SOURCE   Zea mays.
ORGANISM   Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 579)
AUTHORS   Walbot,V.
TITLE   Maize ESTs from various cDNA libraries sequenced at Stanford
               University
JOURNAL   Unpublished (1999)
COMMENT   Contact: Walbot V
               Department of Biological Sciences
               Stanford University
               855 California Ave, Palo Alto, CA 94304, USA
               Tel: 650 723 2227
               Fax: 650 725 8221
               Email: walbot@stanford.edu
               Plate: 683028 row: H column: 09.
FEATURES   Location/Qualifiers
source   1..579
               /organism="Zea mays"
               /cultivar="B73"
               /db_xref="taxon:4577"
               /clone_lib="683 - 14 day immature embryo from Hake lab (HS
)"
               /tissue_type="embryo"
               /dev_stage="14 days after pollination"
               /lab_host="DH10B"
               /note="Organ: embryo; Vector: pBKCMV (Stratagene's Zap
               Express); Site_1: XhoI; Site_2: EcoRI; Directionally
               cloned, 14 day immature embryo library created with
               Stratagene's Zap Express cDNA protocol."
BASE COUNT   119 a   150 c   167 g   143 t
ORIGIN

Query Match   87.6%; Score 18.4; DB 10; Length 579;
Best Local Similarity 95.0%; Pred. No. 5.3e+02;
Matches   19; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
        |||||||||||||||| ||||
Db 107 TCATGGTGATGATAGGAATG 88

```
            then went through one round of normalization to CoT value
            of 5 based on the methods of Marcelo Bento Soares (Genome
            Research 6: 791-806, 1996)."
BASE COUNT      129 a      153 c      166 g      147 t
ORIGIN

Query Match          87.6%;  Score 18.4;  DB 13;  Length 595;
Best Local Similarity  95.0%;  Pred. No. 5.4e+02;
Matches  19;   Conservative  0;   Mismatches  1;   Indels  0;   Gaps  0;

Qy        2 TCATGGTGATGATTGGAATG 21
            ||||||||||||| ||||||
Db      125 TCATGGTGATGATAGGAATG 106

RESULT 13
AW076451/c
LOCUS       AW076451                 610 bp    mRNA    linear  EST 14-OCT-1999
DEFINITION  683018E04.x1 683 - 14 day immature embryo from Hake lab (HS) Zea
            mays cDNA, mRNA sequence.
ACCESSION   AW076451
VERSION     AW076451.1  GI:6031549
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS     Walbot,V.
TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
            University
JOURNAL     Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 683018 row: E column: 04.
FEATURES             Location/Qualifiers
source          1..610
                /organism="Zea mays"
                /cultivar="B73"
                /db_xref="taxon:4577"
                /clone_lib="683 - 14 day immature embryo from Hake lab (HS
                )"
                /tissue_type="embryo"
                /dev_stage="14 days after pollination"
                /lab_host="DH10B"
                /note="Organ: embryo; Vector: pBKCMV (Stratagene's Zap
                Express); Site_1: XhoI; Site_2: EcoRI; Directionally
                cloned, 14 day immature embryo library created with
                Stratagene's Zap Express cDNA protocol."
BASE COUNT     126 a      155 c      180 g      148 t
ORIGIN

Query Match          87.6%;  Score 18.4;  DB 10;  Length 610;
Best Local Similarity  95.0%;  Pred. No. 5.4e+02;
Matches  19;   Conservative  0;   Mismatches  1;   Indels  0;   Gaps  0;

Qy        2 TCATGGTGATGATTGGAATG 21
            ||||||||||||| ||||||
Db      107 TCATGGTGATGATAGGAATG 88

RESULT 14
AV702181
LOCUS       AV702181                 623 bp    mRNA    linear  EST 08-OCT-2000
DEFINITION  AV702181 ADB Homo sapiens cDNA clone ADBDAA11 5', mRNA sequence.
ACCESSION   AV702181
VERSION     AV702181.1  GI:10718511
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 623)
AUTHORS     Peng,Y., Song,H., Huang,Q., Huang,C., Gu,Y., Yang,Y., Xiao
            ,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z.,
            Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
            ,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE       Homo sapiens cDNA ADB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
source          1..623
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="ADBDAA11"
                /clone_lib="ADB"
                /tissue_type="Adrenal gland"
                /dev_stage="Adult"
                /lab_host="SOLR"
                /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                XhoI"
BASE COUNT      171 a      112 c      115 g      217 t      8 others
ORIGIN

Query Match          87.6%;  Score 18.4;  DB 10;  Length 623;
Best Local Similarity  95.0%;  Pred. No. 5.5e+02;
Matches  19;   Conservative  0;   Mismatches  1;   Indels  0;   Gaps  0;

Qy        2 TCATGGTGATGATTGGAATG 21
            ||||||||||||  ||||||
Db      554 TCATGGTGATGAGGAATG 573

RESULT 15
BM351807/c
LOCUS       BM351807                 634 bp    mRNA    linear  EST 16-JAN-2002
DEFINITION  MEST346-G08.T3 ISUM5-RN Zea mays cDNA clone MEST346-G08 3', mRNA
            sequence.
ACCESSION   BM351807
VERSION     BM351807.1  GI:18176874
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 634)
AUTHORS     Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
TITLE       Expressed Sequence Tags from B73 Maize: various stages and tissues
            including seedlings treated with a variety of hormones
JOURNAL     Unpublished (2001)
COMMENT     Contact: Patrick S. Schnable
            Schnable Laboratory
            Iowa State University
            G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
            Tel: 515-294-0975
            Fax: 515-294-2299
            Email: schnable@iastate.edu
            Individual basecall and confidence value were assigned using the
            Phred software.
            (<http://depts.washington.edu/ventures/collabtr/direct/index.htm>b
            rt). Overall sequence quality assessment and vector trimming were
            conducted using the Lucy software (<http://www.tigr.org/softlab/>).
```

Lucy parameters were set to ensure an overall trimmed quality of
97.5% or better without any vector fragments in the chosen
high-quality region of each sequence. Low-quality bases between the
poly-T and the high-quality region were replaced with N's to serve
as spacers.
PCR PRimers
FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES   Location/Qualifiers
source   1..634
                  /organism="Zea mays"
                  /cultivar="B73"
                  /db_xref="taxon:4577"
                  /clone="MEST346-G08"
                  /clone_lib="ISUM5-RN"
                  /tissue_type="mixed"
                  /lab_host="DH10B"
                  /note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
                  Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
                  Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels
                  (3, 5, 10, 15, 20, 25, 30, DAP), Adventious roots (65 DAG
                  ), Tassel (3-39 cm, 53 und 56 DAG), Immature ear (0.2-3.0
                  cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
                  first ear, ear shank, etiolated seedlings, callus,
                  Cycloheximide-treated callus, Anaerobic treated seedlings
                  , NAA (a-Naphthalene acetic acid)-treated seedlings,
                  Kinetin-treated seedlings, ACPC
                  (1-aminocyclopropane-1-carboxylix acid)-treated seedlings,
                  Brassinolide-treated seedlings, ABA (Abscisic acid
)-treated seedlings, GA (Gibberellic acid)-treated
                  seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA
                  molecules were generated as follows. First-strand cDNA was
                  prepared from oligo-dT selected mRNA by priming with a
                  NotI oligo-dT primer (5'
                  AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT). The
                  resulting DNA:RNA hybrid was treated with RNase H and used
                  as a template for DNA PolI-catalyzed second strand
                  synthesis. After the addition of EcoRI adaptors, the
                  ds-cDNAs were digested with NotI and size-selected. The
                  resulting molecules were directionally cloned into the
                  EcoRI and NotI sites of the pT7T3PAC vector. The library
                  then went through one round of normalization to CoT value
                  of 5 based on the methods of Marcelo Bento Soares (Genome
                  Research 6: 791-806, 1996)."
BASE COUNT   128 a   161 c   196 g   149 t
ORIGIN

Query Match   87.6%;   Score 18.4;   DB 13;   Length 634;
Best Local Similarity   95.0%;   Pred. No. 5.5e+02;
Matches   19;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   2 TCATGGTGATGATTGGAATG 21
           ||||||||||||| ||||||
Db   104 TCATGGTGATGATAGGAATG 85

Search completed: June 9, 2003, 10:05:54
Job time : 888 secs

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 12:07:00 ; Search time 39 Seconds
                                    (without alignments)
                                    1423.295 Million cell updates/sec

Title:          US-09-550-163-1_COPY_80_260
Perfect score:  181
Sequence:       1 acttatccaattcacaca.........gattggaatgtcctcttca 181

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       441362 seqs, 153338381 residues

Word size :     0

Total number of hits satisfying chosen parameters:        882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :      Issued Patents_NA:*
           1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
           6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
c 1	9.9	18	1841	1	US-08-570-311-5	Sequence 5, Appli
c 2	9.9	18	1841	1	US-08-353-485-5	Sequence 5, Appli
c 3	9.9	18	2804	1	US-08-446-794A-3	Sequence 3, Appli
c 4	9.4	17	249	1	US-08-378-698-2	Sequence 2, Appli
c 5	9.4	17	249	5	PCT-US96-00728-2	Sequence 2, Appli
c 6	9.4	17	2804	1	US-08-446-794A-3	Sequence 3, Appli
c 7	8.8	16	622	4	US-09-328-111-381	Sequence 381, App
c 8	8.8	16	1626	4	US-09-342-648-7	Sequence 7, Appli
9	8.8	16	14796	4	US-08-975-080-35	Sequence 35, Appl
10	8.8	16	14796	4	US-09-630-706-10	Sequence 10, Appl
11	8.8	16	14796	4	US-09-496-694B-3	Sequence 3, Appli
12	8.3	15	797	4	US-09-221-017B-994	Sequence 994, App
c 14	8.3	15	1854	4	US-09-134-001C-1175	Sequence 1175, Ap
c 15	8.3	15	2069	4	US-08-961-527-174	Sequence 174, App
16	8.3	15	2608	3	US-08-804-439A-12	Sequence 12, Appli
c 17	8.3	15	2608	3	US-08-720-229-12	Sequence 12, Appli
18	8.3	15	3258	2	US-08-729-955A-2	Sequence 2, Appli
c 19	8.3	15	3357	3	US-08-726-214-7	Sequence 7, Appli
c 20	8.3	15	4965	3	US-08-961-527-143	Sequence 143, App
21	8.3	15	12666	4	US-08-961-527-137	Sequence 137, Appli
c 22	8.3	15	13440	3	US-08-961-527-128	Sequence 128, App
c 23	8.3	15	15144	3	US-08-458-434A-6	Sequence 6, Appli
24	8.3	15	34303	2	US-08-735-609-4	Sequence 4, Appli
c 25	8.3	15	34303	3	US-09-315-372-4	Sequence 4, Appli
26	8.3	15	34303	3	US-09-244-752-4	Sequence 4, Appli
27	8.3	15	34303	3	US-09-245-497-4	Sequence 4, Appli
			34303	4	US-08-374-483-6	Sequence 6, Appli
28	8.3	15	34382	2	US-08-374-483-6	Sequence 6, Appli
29	8.3	15	35408	4	US-08-973-334-3	Sequence 3, Appli
30	8.3	15	35408	4	US-09-563-869A-3	Sequence 3, Appli
31	8.3	15	35408	4	US-08-549-489-3	Sequence 3, Appli
32	8.3	15	35935	2	US-08-735-609-1	Sequence 1, Appli
33	8.3	15	35935	2	US-08-735-609-1	Sequence 1, Appli
34	8.3	15	35935	3	US-08-379-452-43	Sequence 43, Appli
35	8.3	15	35935	3	US-09-315-372-1	Sequence 1, Appli
36	8.3	15	35935	3	US-09-244-752-1	Sequence 1, Appli
37	8.3	15	35935	3	US-09-245-497-1	Sequence 1, Appli
38	8.3	15	35935	4	US-09-409-670-43	Sequence 43, Appli
39	8.3	15	35935	1	US-09-562-919-1	Sequence 1, Appli
40	8.3	15	72604	4	US-09-288-992-7	Sequence 7, Appli
41	7.7	14	72604	4	US-09-657-474-7	Sequence 7, Appli
c 42	7.7	14	17	4	US-08-584-040-7504	Sequence 7504, Ap
c 43	7.7	14	17	4	US-08-584-040-7505	Sequence 7505, Ap
44	7.7	14	24	4	US-08-062-368-12	Sequence 12, Appl
c 45	7.7	14	30	4	US-09-593-32-13	Sequence 13, Appl
c 46	7.7	14	30	4	US-09-594-106-13	Sequence 13, Appl
c 47	7.7	14	30	4	US-09-344-300-13	Sequence 13, Appl
c 48	7.7	14	71	1	US-08-464-318-11	Sequence 11, Appl
49	7.7	14	71	1	US-08-464-318-12	Sequence 12, Appl
c 50	7.7	14	71	1	US-08-471-341-11	Sequence 11, Appl
51	7.7	14	71	2	US-08-471-341-12	Sequence 12, Appl
52	7.7	14	71	2	US-08-461-566-11	Sequence 11, Appl
c 53	7.7	14	71	2	US-08-461-566-12	Sequence 12, Appl
54	7.7	14	87	2	US-08-477-527A-126	Sequence 126, App
55	7.7	14	87	2	US-08-481-710-126	Sequence 126, App
56	7.7	14	87	2	PCT-US96-09537-126	Sequence 126, App
57	7.7	14	101	1	US-08-445-133A-12	Sequence 12, Appl
58	7.7	14	101	1	US-08-445-133A-13	Sequence 13, Appl
59	7.7	14	132	1	US-08-445-133A-7	Sequence 7, Appl
60	7.7	14	132	1	US-08-445-133A-9	Sequence 9, Appli
61	7.7	14	132	1	US-08-445-133A-11	Sequence 11, Appl
62	7.7	14	136	1	US-08-445-133A-6	Sequence 6, Appli
63	7.7	14	136	1	US-08-445-133A-8	Sequence 8, Appli
64	7.7	14	136	1	US-08-445-133A-10	Sequence 10, Appl
65	7.7	14	247	4	US-09-344-888A-24	Sequence 24, Appl
66	7.7	14	250	4	US-08-840-466A-25	Sequence 25, Appl
67	7.7	14	250	4	US-09-696-188B-25	Sequence 25, Appl
68	7.7	14	277	4	US-09-060-756-66	Sequence 66, Appl
69	7.7	14	335	2	US-08-675-508-12	Sequence 12, Appl
70	7.7	14	393	3	US-09-188-930-26	Sequence 26, Appl
c 71	7.7	14	401	3	US-08-747-221B-1	Sequence 1, Appli
c 72	7.7	14	401	3	US-08-747-221B-3	Sequence 3, Appli
c 73	7.7	14	401	4	US-09-005-051-1	Sequence 1, Appli
c 74	7.7	14	401	4	US-09-005-051-3	Sequence 3, Appli
75	7.7	14	458	3	US-09-080-855-17	Sequence 17, Appl
76	7.7	14	520	1	US-08-268-348A-7	Sequence 7, Appli
77	7.7	14	520	1	US-08-268-348A-9	Sequence 9, Appli
78	7.7	14	537	2	US-08-176-620A-9	Sequence 9, Appli
79	7.7	14	537	2	US-08-461-985-9	Sequence 9, Appli
80	7.7	14	537	4	US-08-932-787B-17	Sequence 17, Appl
81	7.7	14	537	4	US-08-932-012C-17	Sequence 17, Appl
82	7.7	14	599	1	US-08-888-818C-17	Sequence 17, Appl
83	7.7	14	601	4	US-08-584-226-3	Sequence 3, Appli
84	7.7	14	615	4	US-09-336-536-74	Sequence 74, Appl
85	7.7	14	708	1	US-08-539-304A-1	Sequence 1, Appli
c 86	7.7	14	815	4	US-08-078-090-24	Sequence 24, Appl
c 87	7.7	14	926	3	US-09-221-017B-399	Sequence 399, App
c 88	7.7	14	926	4	US-08-294-871A-67	Sequence 67, Appli
89	7.7	14	957	4	US-08-268-348A-9	Sequence 9, Appli
90	7.7	14	975	1	US-08-876-398A-67	Sequence 67, Appli
91	7.7	14	975	4	US-09-766-055A-1	Sequence 1, Appli
c 92	7.7	14	1000	1	US-09-134-001C-2159	Sequence 2159, Ap
c 93	7.7	14	1000	5	US-08-599-252-96	Sequence 96, Appl
c 94	7.7	14	1001	3	PCT-US96-06352-96	Sequence 96, Appl
95	7.7	14	1001	5	PCT-US96-06583-96	Sequence 96, Appl
c 96	7.7	14	1015	4	US-09-188-930-218	Sequence 218, App
c 97	7.7	14	1041	3	US-09-188-930-30	Sequence 30, Appl
c 98	7.7	14	1047	4	US-09-453-702B-189	Sequence 189, App
99	7.7	14	1065	4	US-09-105-390-47	Sequence 47, Appli
c 100	7.7	14		1	US-08-249-554-1	Sequence 1, Appli

No.			Length	N	Reference	Sequence
c 101	14	7.7	1065	1	US-08-249-555A-1	Sequence 1, Appli
c 102	14	7.7	1065	1	US-08-249-584-1	Sequence 1, Appli
c 103	14	7.7	1065	1	US-08-734-792-1	Sequence 1, Appli
c 104	14	7.7	1065	1	US-08-078-090-1	Sequence 1, Appli
c 105	14	7.7	1065	5	PCT-US95-03628-1	Sequence 1, Appli
c 106	14	7.7	1065	5	PCT-US95-03789-1	Sequence 4, Appli
107	14	7.7	1106	1	US-08-041-648-4	Sequence 6, Appli
108	14	7.7	1106	1	US-08-041-648-6	Sequence 8, Appli
109	14	7.7	1106	1	US-08-041-648-8	Sequence 40, Appl
110	14	7.7	1113	4	US-09-242-690A-40	Sequence 63, Appl
111	14	7.7	1131	4	US-09-105-390-63	Sequence 602, App
112	14	7.7	1149	4	US-09-134-001C-602	Sequence 19, Appl
c 113	14	7.7	1249	3	US-08-755-587-19	Sequence 23, Appl
114	14	7.7	1275	2	US-08-741-437-2	Sequence 2, Appli
115	14	7.7	1275	2	US-09-134-593-2	Sequence 29, Appl
116	14	7.7	1303	1	US-08-265-006-1	Sequence 1, Appli
c 117	14	7.7	1305	6	5212087-1	Patent No. 5212087
c 118	14	7.7	1312	3	US-08-961-083-69	Sequence 69, Appl
119	14	7.7	1559	3	US-08-294-871A-33	Sequence 33, Appl
120	14	7.7	1559	1	US-08-294-871A-35	Sequence 35, Appl
121	14	7.7	1559	1	US-08-294-871A-37	Sequence 37, Appl
122	14	7.7	1559	1	US-08-294-871A-39	Sequence 39, Appl
123	14	7.7	1559	1	US-08-294-871A-41	Sequence 41, Appl
124	14	7.7	1559	1	US-08-294-871A-43	Sequence 43, Appl
125	14	7.7	1559	1	US-08-294-871A-45	Sequence 45, Appl
126	14	7.7	1559	1	US-08-294-871A-53	Sequence 53, Appl
127	14	7.7	1559	1	US-08-294-871A-57	Sequence 57, Appl
128	14	7.7	1559	3	US-08-876-398A-33	Sequence 33, Appl
129	14	7.7	1559	3	US-08-876-398A-35	Sequence 35, Appl
130	14	7.7	1559	3	US-08-876-398A-37	Sequence 37, Appl
131	14	7.7	1559	3	US-08-876-398A-39	Sequence 39, Appl
132	14	7.7	1559	3	US-08-876-398A-41	Sequence 41, Appl
133	14	7.7	1559	3	US-08-876-398A-43	Sequence 43, Appl
134	14	7.7	1559	3	US-08-876-398A-45	Sequence 45, Appl
135	14	7.7	1559	3	US-08-876-398A-53	Sequence 53, Appl
136	14	7.7	1559	3	US-08-876-398A-57	Sequence 57, Appl
137	14	7.7	1584	4	US-09-318-448-31	Sequence 31, Appl
138	14	7.7	1588	4	US-08-617-785-7	Sequence 7, Appli
139	14	7.7	1695	2	US-08-878-957-31	Sequence 31, Appl
140	14	7.7	1710	3	US-08-747-221B-27	Sequence 27, Appl
141	14	7.7	1710	3	US-08-747-221B-33	Sequence 33, Appl
142	14	7.7	1710	4	US-09-005-051-27	Sequence 27, Appl
143	14	7.7	1710	4	US-09-005-051-33	Sequence 33, Appl
144	14	7.7	1740	2	US-08-796-488-1	Sequence 1, Appli
145	14	7.7	1740	4	US-09-243-914-1	Sequence 1, Appli
c 146	14	7.7	1777	4	US-09-058-260-25	Sequence 25, Appl
147	14	7.7	1785	1	US-07-917-111-1	Sequence 1, Appli
148	14	7.7	1785	1	US-08-479-638-1	Sequence 1, Appli
149	14	7.7	1785	1	US-08-294-871A-1	Sequence 1, Appli
150	14	7.7	1785	1	US-08-294-871A-3	Sequence 3, Appli
151	14	7.7	1785	1	US-08-294-871A-5	Sequence 5, Appli
152	14	7.7	1785	1	US-08-294-871A-7	Sequence 7, Appli
153	14	7.7	1785	1	US-08-294-871A-9	Sequence 9, Appli
154	14	7.7	1785	1	US-08-294-871A-11	Sequence 11, Appl
155	14	7.7	1785	1	US-08-294-871A-13	Sequence 13, Appl
156	14	7.7	1785	1	US-08-294-871A-15	Sequence 15, Appl
157	14	7.7	1785	1	US-08-294-871A-17	Sequence 17, Appl
158	14	7.7	1785	1	US-08-294-871A-19	Sequence 19, Appl
159	14	7.7	1785	1	US-08-294-871A-21	Sequence 21, Appl
160	14	7.7	1785	1	US-08-294-871A-23	Sequence 23, Appl
161	14	7.7	1785	1	US-08-294-871A-25	Sequence 25, Appl
162	14	7.7	1785	1	US-08-294-871A-27	Sequence 27, Appl
163	14	7.7	1785	1	US-08-294-871A-29	Sequence 29, Appl
164	14	7.7	1785	1	US-08-294-871A-31	Sequence 31, Appl
165	14	7.7	1785	1	US-08-294-871A-47	Sequence 47, Appl
166	14	7.7	1785	1	US-08-294-871A-49	Sequence 49, Appl
167	14	7.7	1785	1	US-08-294-871A-51	Sequence 51, Appl
168	14	7.7	1785	1	US-08-294-871A-55	Sequence 55, Appl
169	14	7.7	1785	1	US-08-294-871A-59	Sequence 59, Appl
170	14	7.7	1785	1	US-08-294-871A-61	Sequence 61, Appl
171	14	7.7	1785	1	US-08-294-871A-63	Sequence 63, Appl
172	14	7.7	1785	1	US-08-294-871A-65	Sequence 65, Appl
173	14	7.7	1785	3	US-08-747-221B-34	Sequence 34, Appl
c 174	14	7.7	1785	3	US-08-747-221B-35	Sequence 35, Appl
175	14	7.7	1785	3	US-08-876-398A-1	Sequence 1, Appli
176	14	7.7	1785	3	US-08-876-398A-3	Sequence 3, Appli
177	14	7.7	1785	3	US-08-876-398A-5	Sequence 5, Appli
178	14	7.7	1785	3	US-08-876-398A-7	Sequence 7, Appli
179	14	7.7	1785	3	US-08-876-398A-9	Sequence 9, Appli
180	14	7.7	1785	3	US-08-876-398A-11	Sequence 11, Appl
181	14	7.7	1785	3	US-08-876-398A-13	Sequence 13, Appl
182	14	7.7	1785	3	US-08-876-398A-15	Sequence 15, Appl
183	14	7.7	1785	3	US-08-876-398A-17	Sequence 17, Appl
184	14	7.7	1785	3	US-08-876-398A-19	Sequence 19, Appl
185	14	7.7	1785	3	US-08-876-398A-21	Sequence 21, Appl
186	14	7.7	1785	3	US-08-876-398A-23	Sequence 23, Appl
187	14	7.7	1785	3	US-08-876-398A-25	Sequence 25, Appl
188	14	7.7	1785	3	US-08-876-398A-27	Sequence 27, Appl
189	14	7.7	1785	3	US-08-876-398A-29	Sequence 29, Appl
190	14	7.7	1785	3	US-08-876-398A-31	Sequence 31, Appl
191	14	7.7	1785	3	US-08-876-398A-47	Sequence 47, Appl
192	14	7.7	1785	3	US-08-876-398A-49	Sequence 49, Appl
193	14	7.7	1785	3	US-08-876-398A-51	Sequence 51, Appl
194	14	7.7	1785	3	US-08-876-398A-55	Sequence 55, Appl
195	14	7.7	1785	3	US-08-876-398A-59	Sequence 59, Appl
196	14	7.7	1785	3	US-08-876-398A-61	Sequence 61, Appl
197	14	7.7	1785	3	US-08-876-398A-63	Sequence 63, Appl
198	14	7.7	1785	3	US-08-876-398A-65	Sequence 65, Appl
199	14	7.7	1785	4	US-09-005-051-34	Sequence 34, Appl
c 200	14	7.7	1785	4	US-09-005-051-35	Sequence 35, Appl
c 201	14	7.7	1788	3	US-08-747-221B-28	Sequence 28, Appl
c 202	14	7.7	1788	3	US-08-747-221B-29	Sequence 29, Appl
c 203	14	7.7	1788	4	US-09-005-051-28	Sequence 28, Appl
c 204	14	7.7	1788	4	US-09-005-051-29	Sequence 29, Appl
c 205	14	7.7	1810	4	US-08-714-918-32	Sequence 32, Appl
c 206	14	7.7	1810	4	US-09-265-315-32	Sequence 32, Appl
c 207	14	7.7	1810	4	US-09-265-315-32	Sequence 32, Appl
c 208	14	7.7	1810	4	US-09-266-417-32	Sequence 32, Appl
c 209	14	7.7	1816	1	US-07-862-021B-13	Sequence 13, Appl
c 210	14	7.7	1816	1	US-08-313-288B-13	Sequence 13, Appl
c 211	14	7.7	1816	5	PCT-US93-03164-13	Sequence 13, Appl
212	14	7.7	1826	2	US-08-808-931-15	Sequence 15, Appl
213	14	7.7	1826	3	US-08-808-323-15	Sequence 15, Appl
214	14	7.7	1826	3	US-09-050-603A-15	Sequence 15, Appl
215	14	7.7	1826	3	US-09-102-420B-15	Sequence 15, Appl
216	14	7.7	1826	3	US-09-497-698-15	Sequence 15, Appl
217	14	7.7	1841	1	US-08-570-311-5	Sequence 5, Appli
218	14	7.7	1841	2	US-08-353-485-5	Sequence 5, Appli
219	14	7.7	1894	1	US-08-935-450-7	Sequence 7, Appli
220	14	7.7	1915	1	US-08-496-855A-5	Sequence 5, Appli
221	14	7.7	1915	2	US-08-466-589-11	Sequence 11, Appl
222	14	7.7	1915	2	US-08-700-036-11	Sequence 11, Appl
223	14	7.7	1915	2	US-08-467-574-11	Sequence 11, Appl
224	14	7.7	1915	4	US-08-217-345-11	Sequence 11, Appl
225	14	7.7	1924	4	US-08-487-596-17	Sequence 17, Appl
226	14	7.7	1941	4	US-09-058-260-29	Sequence 29, Appl
227	14	7.7	2001	4	US-09-620-412C-316	Sequence 316, App
228	14	7.7	2086	4	US-09-134-001C-1902	Sequence 1902, Ap
229	14	7.7	2086	2	US-08-557-128-5	Sequence 5, Appli
230	14	7.7	2122	4	US-09-242-690A-37	Sequence 37, Appl
c 231	14	7.7	2184	4	US-09-058-260-8	Sequence 8, Appli
232	14	7.7	2230	1	US-09-105-390-8	Sequence 9, Appli
233	14	7.7	2315	1	US-08-200-512-1	Sequence 1, Appli
234	14	7.7	2345	4	US-09-058-260-9	Sequence 9, Appli
c 235	14	7.7	2405	4	US-09-058-260-15	Sequence 15, Appl
c 236	14	7.7	2431	4	US-09-058-260-27	Sequence 27, Appl
237	14	7.7	2431	2	US-08-714-918-15	Sequence 15, Appl
238	14	7.7	2431	3	US-09-265-315-15	Sequence 15, Appl
239	14	7.7	2431	3	US-09-265-315-15	Sequence 15, Appl
240	14	7.7	2469	3	US-09-266-417B-351	Sequence 351, App
c 241	14	7.7	2643	3	US-09-221-017B-351	Sequence 51, Appl
c 242	14	7.7	2643	2	US-08-858-207A-51	Sequence 11, Appl
243	14	7.7	2643	2	US-08-781-802-11	Sequence 30, Appl
244	14	7.7	2801	3	US-09-058-260-11	Sequence 32, Appl
245	14	7.7	2801	3	US-08-747-221B-30	Sequence 30, Appl
c 246	14	7.7	2801	3	US-08-747-221B-32	Sequence 32, Appl

Left column (rows 247–319)

#	Count	Tm	Len	Accession	N	Sequence ref
247	2801	7.7	14	US-09-005-051-30	4	Sequence 30, Appl
248	2801	7.7	14	US-09-005-051-32	4	Sequence 32, Appl
249	2836	7.7	14	US-08-747-221B-24	3	Sequence 26, Appl
250	2836	7.7	14	US-08-747-221B-26	4	Sequence 24, Appl
251	2836	7.7	14	US-09-005-051-24	4	Sequence 24, Appl
252	2836	7.7	14	US-09-005-051-26	4	Sequence 26, Appl
253	2885	7.7	14	US-09-232-200-36	4	Sequence 36, Appl
254	2885	7.7	14	US-09-232-200-56	4	Sequence 56, Appl
255	2885	7.7	14	US-09-232-197-36	4	Sequence 36, Appl
256	2885	7.7	14	US-09-232-197-56	4	Sequence 56, Appl
257	2885	7.7	14	US-09-232-201-36	4	Sequence 36, Appl
258	2885	7.7	14	US-09-232-201-56	4	Sequence 56, Appl
259	3030	7.7	14	US-08-680-326-24	4	Sequence 24, Appl
260	3513	7.7	14	US-08-781-860-1	2	Sequence 1, Appli
261	3545	7.7	14	US-08-781-802-9	4	Sequence 9, Appli
262	3558	7.7	14	US-09-134-001C-2439	2	Sequence 2439, Ap
263	3925	7.7	14	US-08-041-648-1	4	Sequence 1, Appli
264	3977	7.7	14	US-06-217-529-1	1	Sequence 1, Appli
265	3977	7.7	14	US-08-781-802-1	4	Sequence 1, Appli
266	4315	7.7	14	US-08-694-078-1	2	Sequence 1, Appli
267	4491	7.7	14	US-09-837-863-23	4	Sequence 23, Appl
268	4755	7.7	14	US-09-837-863-24	4	Sequence 24, Appl
269	4940	7.7	14	US-08-484-105-1	4	Sequence 1, Appli
270	4940	7.7	14	US-08-484-106-1	1	Sequence 1, Appli
271	5095	7.7	14	US-08-092-817-3	4	Sequence 3, Appli
272	5095	7.7	14	US-08-485-128-3	1	Sequence 3, Appli
273	5156	7.7	14	US-09-091-432-3	2	Sequence 3, Appli
274	5186	7.7	14	US-08-596-366-1	4	Sequence 1, Appli
275	5186	7.7	14	US-08-967-104-1	1	Sequence 1, Appli
276	5253	7.7	14	US-08-714-918-19	4	Sequence 19, Appl
277	5253	7.7	14	US-09-265-315-19	4	Sequence 19, Appl
278	5253	7.7	14	US-09-265-315-19	4	Sequence 19, Appl
279	5253	7.7	14	US-09-266-417-19	4	Sequence 19, Appl
280	5883	7.7	14	US-08-404-445-1	4	Sequence 1, Appli
281	6519	7.7	14	US-08-233-008A-7	4	Sequence 7, Appli
282	6873	7.7	14	US-09-131-028A-1	4	Sequence 1, Appli
283	6873	7.7	14	US-08-250-848-8	2	Sequence 8, Appli
284	7559	7.7	14	US-09-131-028A-8	4	Sequence 1, Appli
285	8119	7.7	14	US-08-961-527-74	1	Sequence 74, Appl
286	8119	7.7	14	US-08-398-028B-1	2	Sequence 87, Appl
287	8119	7.7	14	US-08-262-773-206	1	Sequence 206, App
288	8119	7.7	14	US-08-504-265B-1	2	Sequence 9, Appli
289	8430	7.7	14	US-09-131-028A-6	4	Sequence 210, App
290	8430	7.7	14	US-09-131-028A-10	1	Sequence 48, Appl
291	10607	7.7	14	US-08-078-090-3	4	Sequence 22, Appl
292	16535	7.7	14	US-08-961-527-22	4	Sequence 22, Appl
293	18436	7.7	14	US-08-961-527-87	4	Sequence 140, App
294	20137	7.7	14	US-09-262-773-206	4	Sequence 7, Appli
295	20138	7.7	14	US-09-262-773-9	2	Sequence 15, Appl
296	23071	7.7	14	US-09-262-773-210	4	Sequence 15, Appl
297	25002	7.7	14	US-08-961-527-48	4	Sequence 6, Appli
298	28171	7.7	14	US-08-961-527-22	4	Sequence 22, Appl
299	28171	7.7	14	US-08-961-527-140	4	Sequence 22, Appl
300	28882	7.7	14	US-08-814-095-7	3	Sequence 140, App
301	35060	7.7	14	US-08-846-111D-15	4	Sequence 7, Appli
302	40352	7.7	14	US-09-443-077-15	4	Sequence 15, Appl
303	40352	7.7	14	US-09-146-053-6	4	Sequence 15, Appl
304	45546	7.7	14	US-09-676-610B-24	4	Sequence 6, Appli
305	169998	7.7	14	US-09-103-840A-2	4	Sequence 24, Appl
306	4403765	7.7	14	US-09-103-840A-1	4	Sequence 2, Appli
307	4403765	7.7	14	US-09-103-840A-1	4	Sequence 1, Appli
308	4411529	7.7	14	US-08-334-847-75	4	Sequence 1, Appli
309	4411529	7.7	14	US-09-112-096-13	1	Sequence 75, Appl
310	15	7.2	13	US-08-334-847-75	1	Sequence 75, Appl
311	16	7.2	13	US-09-112-096-13	1	Sequence 13, Appl
312	17	7.2	13	US-08-117-907-13	1	Sequence 13, Appl
313	17	7.2	13	US-08-453-956-1	2	Sequence 1, Appli
314	17	7.2	13	US-08-086-631-1	1	Sequence 1, Appli
315	17	7.2	13	US-08-452-930-1	1	Sequence 1, Appli
316	17	7.2	13	US-09-211-631-18	2	Sequence 18, Appl
317	17	7.2	13	US-09-081-180-18	3	Sequence 18, Appl
318	17	7.2	13	US-08-453-040-13	3	Sequence 18, Appl
319	17	7.2	13	US-09-040-786-18	3	Sequence 18, Appl

Right column (rows 320–392)

#	Count	Tm	Len	N	Accession	Sequence ref
320	17	7.2	13	3	US-09-265-628-18	Sequence 18, Appl
321	17	7.2	13	4	US-09-001-141-16	Sequence 16, Appl
322	17	7.2	13	4	US-09-140-804-11	Sequence 11, Appl
323	17	7.2	13	4	US-09-522-217-76	Sequence 76, Appl
324	17	7.2	13	4	US-09-653-403-19	Sequence 19, Appl
325	17	7.2	13	4	US-09-294-531B-14	Sequence 14, Appl
326	17	7.2	13	4	US-09-434-408-11	Sequence 11, Appl
327	17	7.2	13	5	US-10-013-784-19	Sequence 19, Appl
328	18	7.2	13	1	PCT-US93-08174-1	Sequence 1, Appli
329	18	7.2	13	2	US-08-861-775-6	Sequence 6, Appli
330	18	7.2	13	1	US-08-833-887-6	Sequence 6, Appli
331	18	7.2	13	2	US-08-823-603-6	Sequence 6, Appli
332	18	7.2	13	3	US-09-256-465-43	Sequence 43, Appl
333	18	7.2	13	3	US-09-257-710-6	Sequence 6, Appli
334	18	7.2	13	4	US-08-621-700-11	Sequence 11, Appl
335	18	7.2	13	5	US-08-971-099-9	Sequence 9, Appli
336	18	7.2	13	4	US-08-929-940-11	Sequence 11, Appl
337	18	7.2	13	5	PCT-US95-03940-11	Sequence 11, Appl
338	18	7.2	13	4	US-09-509-565-2	Sequence 2, Appli
339	19	7.2	13	1	US-08-117-361C-9	Sequence 9, Appli
340	20	7.2	13	3	US-09-023-221A-17	Sequence 17, Appl
341	20	7.2	13	2	US-08-834-901-2	Sequence 2, Appli
342	20	7.2	13	4	US-09-282-352A-17	Sequence 17, Appl
343	20	7.2	13	4	US-09-287-781A-10	Sequence 10, Appl
344	20	7.2	13	4	US-09-414-276-19	Sequence 19, Appl
345	22	7.2	13	1	US-07-638-512-4	Sequence 4, Appli
346	22	7.2	13	2	US-08-066-325-41	Sequence 41, Appl
347	22	7.2	13	2	US-08-713-298B-7	Sequence 7, Appli
348	22	7.2	13	2	US-08-766-858A-2	Sequence 2, Appli
349	22	7.2	13	2	US-08-870-180B-7	Sequence 7, Appli
350	22	7.2	13	4	US-09-226-529-7	Sequence 13, Appl
351	22	7.2	13	4	US-09-254-352B-13	Sequence 13, Appl
352	23	7.2	13	4	US-08-448-418-68	Sequence 68, Appl
353	23	7.2	13	2	US-08-652-816A-35	Sequence 35, Appl
354	23	7.2	13	2	US-08-350-260A-114	Sequence 114, App
355	23	7.2	13	1	US-09-018-576-14	Sequence 14, Appl
356	23	7.2	13	1	US-09-958-201-20	Sequence 20, Appl
357	23	7.2	13	1	US-09-248-137-14	Sequence 14, Appl
358	23	7.2	13	3	US-09-184-658-39	Sequence 39, Appl
359	23	7.2	13	1	US-08-918-144-54	Sequence 54, Appl
360	23	7.2	13	2	US-09-079-029-12	Sequence 12, Appl
361	23	7.2	13	1	US-09-363-939A-13	Sequence 13, Appl
362	24	7.2	13	1	US-09-343-785-5	Sequence 5, Appli
363	24	7.2	13	1	US-08-049-264C-50	Sequence 50, Appl
364	24	7.2	13	1	US-07-910-867B-17	Sequence 17, Appl
365	24	7.2	13	1	US-08-488-212A-12	Sequence 12, Appl
366	24	7.2	13	1	US-08-387-845-7	Sequence 7, Appli
367	24	7.2	13	1	US-08-346-613-17	Sequence 17, Appl
368	24	7.2	13	1	US-08-476-562-50	Sequence 50, Appl
369	24	7.2	13	1	US-08-400-580A-14	Sequence 14, Appl
370	24	7.2	13	1	US-08-479-723A-50	Sequence 50, Appl
371	24	7.2	13	1	US-08-485-692-18	Sequence 18, Appl
372	24	7.2	13	1	US-08-419-519-18	Sequence 18, Appl
373	24	7.2	13	1	US-08-462-221-5	Sequence 5, Appli
374	24	7.2	13	2	US-08-818-514-16	Sequence 16, Appl
375	24	7.2	13	2	US-08-484-956-51	Sequence 51, Appl
376	24	7.2	13	2	US-08-757-653-51	Sequence 51, Appl
377	24	7.2	13	2	US-08-715-941-5	Sequence 5, Appli
378	24	7.2	13	2	US-08-482-182-49	Sequence 49, Appl
379	24	7.2	13	2	US-08-482-182-56	Sequence 56, Appl
380	24	7.2	13	1	US-08-465-380-101	Sequence 101, App
381	24	7.2	13	2	US-08-480-478-14	Sequence 14, Appl
382	24	7.2	13	1	US-08-486-397-101	Sequence 101, App
383	24	7.2	13	1	US-08-486-399-101	Sequence 101, App
384	24	7.2	13	1	US-08-461-965-101	Sequence 101, App
385	24	7.2	13	2	US-08-875-154-34	Sequence 34, Appl
386	24	7.2	13	2	US-08-320-306-12	Sequence 12, Appl
387	24	7.2	13	2	US-08-485-692-18	Sequence 12, Appl
388	24	7.2	13	2	US-08-488-209B-12	Sequence 5, Appli
389	24	7.2	13	2	US-08-408-011-12	Sequence 12, Appl
390	24	7.2	13	2	US-08-778-275-7	Sequence 7, Appli
391	24	7.2	13	2	US-08-964-725-8	Sequence 8, Appli
392	24	7.2	13	2	US-08-326-110A-14	Sequence 6, Appli

Left column

No.		len	ver	col	sub	Accession / Patent	Sequence
393		13	7.2	24	2	US-08-634-641-101	Sequence 101, App
394		13	7.2	24	2	US-08-812-412-4	Sequence 4, Appl
395		13	7.2	24	2	US-08-589-939-69	Sequence 69, Appl
396		13	7.2	24	3	US-09-249-471-101	Sequence 101, App
397		13	7.2	24	3	US-09-249-472-101	Sequence 101, App
398		13	7.2	24	3	US-08-946-458-5	Sequence 5, Appl
399		13	7.2	24	3	US-08-777-147-3	Sequence 3, Appli
400		13	7.2	24	3	US-08-867-352-8	Sequence 8, Appl
401		13	7.2	24	3	US-08-559-205-11	Sequence 11, Appl
402		13	7.2	24	3	US-09-249-451-101	Sequence 101, App
403		13	7.2	24	3	US-08-809-455-101	Sequence 101, App
404		13	7.2	24	3	US-09-249-461-101	Sequence 101, App
405		13	7.2	24	3	US-08-758-662-2	Sequence 2, Appl
406		13	7.2	24	3	US-09-249-448-101	Sequence 101, App
407		13	7.2	24	3	US-09-330-095-14	Sequence 14, Appl
408		13	7.2	24	3	US-09-071-710-19	Sequence 19, Appl
409		13	7.2	24	3	US-09-115-934A-16	Sequence 16, Appl
410		13	7.2	24	4	US-09-121-286-9	Sequence 9, Appli
411		13	7.2	24	4	US-09-287-936-4	Sequence 4, Appl
412		13	7.2	24	4	US-08-912-276-7	Sequence 7, Appl
413		13	7.2	24	4	US-08-534-799-6	Sequence 6, Appli
414		13	7.2	24	4	US-09-180-271-7	Sequence 7, Appl
415		13	7.2	24	4	US-08-844-045C-9	Sequence 9, Appl
416		13	7.2	24	4	US-08-944-483-11	Sequence 11, Appl
417		13	7.2	24	4	US-09-01-870-5	Sequence 5, Appl
418		13	7.2	24	4	US-09-525-397-19	Sequence 19, Appl
419		13	7.2	24	4	US-09-566-876-7	Sequence 7, Appl
420		13	7.2	24	4	US-09-049-698-21	Sequence 21, Appl
421		13	7.2	24	4	US-08-520-946-51	Sequence 51, Appl
422		13	7.2	24	4	US-09-065-383-13	Sequence 13, Appl
423		13	7.2	24	4	US-09-560-419-19	Sequence 19, Appl
424		13	7.2	24	4	US-09-276-600-4	Sequence 4, Appl
425		13	7.2	24	5	PCT-US91-02626-1	Sequence 1, Appl
426		13	7.2	24	5	PCT-US94-04410-50	Sequence 50, Appl
427		13	7.2	24	5	PCT-US95-14418-7	Sequence 7, Appl
428		13	7.2	24	6	PCT-US95-15327-7	Sequence 7, Appli
429		13	7.2	24	6	5223414-2	Patent No. 5223414
430		13	7.2	24	6	5536598-11	Patent No. 5536598
431		13	7.2	25	1	US-08-446-739-9	Sequence 9, Appl
432		13	7.2	25	3	US-08-973-965-6	Sequence 6, Appl
433		13	7.2	25	3	US-09-165-042-20	Sequence 20, Appl
434	c	13	7.2	28	2	US-08-666-473-94	Sequence 94, Appl
435		13	7.2	28	2	US-08-288-663A-7	Sequence 7, Appl
436		13	7.2	29	3	US-08-484-661A-46	Sequence 46, Appl
437		13	7.2	29	3	US-08-656-664-46	Sequence 46, Appl
438		13	7.2	29	5	PCT-US96-09641-46	Sequence 46, Appl
439		13	7.2	30	2	US-09-503-391-17	Sequence 17, Appl
440		13	7.2	30	4	US-09-443-501A-3	Sequence 3, Appl
441		13	7.2	30	4	US-09-560-367A-1	Sequence 1, Appl
442	c	13	7.2	31	4	US-08-981-030-9	Sequence 9, Appl
443	c	13	7.2	31	4	US-07-689-008-12	Sequence 12, Appl
444	c	13	7.2	32	2	US-09-388-917-7	Sequence 7, Appl
445	c	13	7.2	32	4	US-09-079-955-8	Sequence 8, Appl
446	c	13	7.2	33	3	US-08-484-661A-45	Sequence 45, Appl
447	c	13	7.2	33	3	US-08-656-664-45	Sequence 45, Appl
448	c	13	7.2	33	5	PCT-US96-09641-45	Sequence 45, Appl
449	c	13	7.2	34	2	US-09-507-323B-5	Sequence 5, Appl
450	c	13	7.2	40	2	US-08-425-684-80	Sequence 80, Appl
451	c	13	7.2	40	2	US-08-675-502-80	Sequence 80, Appl
452	c	13	7.2	40	2	US-09-086-726-10	Sequence 10, Appl
453	c	13	7.2	42	1	US-08-616-113-6	Sequence 6, Appl
454	c	13	7.2	42	1	US-08-802-995-6	Sequence 6, Appl
455	c	13	7.2	42	1	US-09-129-740-11	Sequence 11, Appl
456	c	13	7.2	42	4	US-09-568-527-11	Sequence 11, Appl
457	c	13	7.2	44	1	US-08-181-629A-1	Sequence 1, Appl
458	c	13	7.2	45	2	US-08-532-795-12	Sequence 12, Appl
459	c	13	7.2	47	5	PCT-US92-10421-8	Sequence 8, Appl
460	c	13	7.2	48	1	US-07-971-819A-16	Sequence 16, Appl
461	c	13	7.2	48	1	US-07-971-819A-18	Sequence 18, Appl
462	c	13	7.2	48	1	US-08-475-331-18	Sequence 18, Appl
463	c	13	7.2	48	5	PCT-US92-10421-7	Sequence 7, Appl
464	c	13	7.2	49	1	US-08-327-525A-19	Sequence 15, Appl
465	c	13	7.2	51	1	US-08-384-490-15	Sequence 15, Appl

Right column

No.		len	ver	col	sub	Accession / Patent	Sequence
466	c	13	7.2	51	1	US-08-384-490-17	Sequence 17, Appl
467	c	13	7.2	51	1	US-08-459-383-15	Sequence 15, Appl
468	c	13	7.2	51	1	US-08-459-383-17	Sequence 17, Appl
469		13	7.2	54	5	PCT-US92-10421-10	Sequence 10, Appl
470		13	7.2	55	2	US-08-594-469-3	Sequence 3, Appli
471		13	7.2	55	4	US-08-906-957-3	Sequence 1, Appli
472	c	13	7.2	59	4	US-09-449-059-1	Sequence 2, Appli
473		13	7.2	59	4	US-09-449-059-2	Sequence 2, Appl
474	c	13	7.2	59	4	US-09-721-801-1	Sequence 15, Appl
475		13	7.2	60	1	US-09-721-801-2	Sequence 19, Appl
476	c	13	7.2	60	1	US-08-247-475-15	Sequence 15, Appl
477	c	13	7.2	60	1	US-08-479-650-15	Sequence 7, Appl
478	c	13	7.2	60	1	US-08-191-866D-19	Sequence 21, Appl
479	c	13	7.2	60	2	US-08-674-169-15	Sequence 50, Appl
480	c	13	7.2	60	4	US-08-185-949B-19	Sequence 25, Appl
481		13	7.2	62	4	US-08-050-482A-21	Sequence 25, Appl
482		13	7.2	63	5	PCT-US95-14418-50	Sequence 27, Appl
483		13	7.2	70	5	PCT-US95-15327-50	Sequence 7, Appl
484		13	7.2	70	5	US-07-920-519-25	Sequence 51, Appl
485		13	7.2	74	1	US-08-314-586-25	Sequence 51, Appl
486		13	7.2	74	3	US-08-115-753-27	Sequence 26, Appl
487		13	7.2	77	1	US-07-910-528-7	Sequence 31, Appl
488		13	7.2	77	1	US-08-348-972-7	Sequence 19, Appl
489		13	7.2	81	5	PCT-US95-14418-51	Sequence 20, Appl
490	c	13	7.2	81	5	PCT-US95-15327-51	Sequence 9, Appl
491	c	13	7.2	83	4	US-09-332-319-26	Sequence 5, Appl
492		13	7.2	85	2	US-08-252-057-31	Sequence 21, Appl
493	c	13	7.2	91	4	US-08-050-482A-19	Sequence 5, Appl
494		13	7.2	91	4	US-08-050-482A-20	Sequence 25, Appl
495	c	13	7.2	95	1	US-08-153-071-2	Sequence 25, Appl
496		13	7.2	95	2	US-08-609-271-9	Sequence 10, Appl
497		13	7.2	95	2	US-08-438-511-2	Sequence 6, Appl
498		13	7.2	95	2	US-08-487-431-5	Sequence 10, Appl
499		13	7.2	95	3	US-09-235-246-21	Sequence 6, Appl
500		13	7.2	95	3	US-08-188-374-9	Sequence 28, Appl
501		13	7.2	95	3	US-08-973-629-5	Sequence 4, Appl
502		13	7.2	96	1	US-08-361-708-25	Sequence 3, Appl
503	c	13	7.2	96	1	US-08-153-071-3	Sequence 3, Appl
504	c	13	7.2	96	2	US-08-536-777-25	Sequence 4, Appl
505	c	13	7.2	96	2	US-08-609-271-10	Sequence 24, Appl
506	c	13	7.2	96	2	US-08-438-511-3	Sequence 39, Appl
507	c	13	7.2	96	3	US-08-487-431-6	Sequence 30, Appl
508	c	13	7.2	96	3	US-08-188-374-10	Sequence 3, Appl
509	c	13	7.2	97	3	US-08-973-629-6	Sequence 22, Appl
510		13	7.2	98	3	US-09-106-638-6	Sequence 10, Appl
511		13	7.2	98	4	US-08-210-222-28	Sequence 28, Appl
512	c	13	7.2	99	4	US-09-721-801-4	Sequence 4, Appl
513		13	7.2	99	4	US-09-721-801-4	Sequence 3, Appl
514	c	13	7.2	99	4	US-09-449-059-3	Sequence 3, Appl
515	c	13	7.2	99	4	US-09-449-059-4	Sequence 4, Appl
516	c	13	7.2	99	4	US-09-721-801-3	Sequence 3, Appl
517		13	7.2	99	4	US-09-721-801-3	Sequence 3, Appl
518	c	13	7.2	100	3	US-09-070-842A-4	Sequence 4, Appl
519		13	7.2	102	4	US-09-332-319-24	Sequence 24, Appl
520		13	7.2	104	4	US-08-353-372A-39	Sequence 39, Appl
521		13	7.2	104	4	US-08-057-430A-30	Sequence 30, Appl
522		13	7.2	109	1	US-07-640-476-3	Sequence 3, Appli
523		13	7.2	114	2	US-07-666-719-22	Sequence 22, Appl
524		13	7.2	114	6	5191064-2	Patent No. 5191064
525		13	7.2	122	3	US-08-411-768B-19	Sequence 19, Appl
526		13	7.2	124	3	US-08-229-444B-3	Sequence 3, Appl
527		13	7.2	125	3	US-08-411-768B-17	Sequence 17, Appl
528		13	7.2	126	3	US-08-411-768B-18	Sequence 18, Appl
529		13	7.2	128	2	US-09-332-319-28	Sequence 28, Appl
530		13	7.2	130	1	US-08-343-785-6	Sequence 6, Appl
531		13	7.2	130	2	US-08-462-221-6	Sequence 6, Appl
532		13	7.2	130	2	US-08-946-458-6	Sequence 6, Appli
533		13	7.2	140	6	5252476-9	Patent No. 5252476
534	c	13	7.2	142	4	US-09-025-769B-266	Sequence 266, App
535		13	7.2	157	5	PCT-US93-06251-61	Sequence 61, Appl
536		13	7.2	160	1	US-08-327-525A-19	Sequence 19, Appl
537		13	7.2	160	1	US-08-327-525A-21	Sequence 21, Appl

	SEQ			Pos	App	Patent No.	Reference
	539	13	7.2	160	1	US-08-327-525A-22	Sequence 22, Appl
	540	13	7.2	160	1	US-08-327-525A-23	Sequence 25, Appl
	541	13	7.2	160	1	US-08-327-525A-25	Sequence 26, Appl
	542	13	7.2	160	1	US-08-327-525A-26	Sequence 27, Appl
	543	13	7.2	160	1	US-08-327-525A-27	Sequence 19, Appl
	544	13	7.2	160	2	US-08-531-137B-19	Sequence 21, Appl
	545	13	7.2	160	2	US-08-531-137B-21	Sequence 23, Appl
	546	13	7.2	160	2	US-08-531-137B-23	Sequence 25, Appl
	547	13	7.2	160	2	US-08-531-137B-25	Sequence 26, Appl
	548	13	7.2	160	2	US-08-531-137B-26	Sequence 27, Appl
	549	13	7.2	160	2	US-08-531-137B-27	Sequence 19, Appl
	550	13	7.2	160	4	US-09-158-765-19	Sequence 21, Appl
	551	13	7.2	160	4	US-09-158-765-21	Sequence 22, Appl
	552	13	7.2	160	4	US-09-158-765-22	Sequence 23, Appl
	553	13	7.2	160	4	US-09-158-765-23	Sequence 26, Appl
	554	13	7.2	160	4	US-09-158-765-25	Sequence 27, Appl
	555	13	7.2	160	4	US-09-158-765-26	Sequence 15, Appl
	556	13	7.2	160	4	US-09-158-765-27	Sequence 22, Appl
	557	13	7.2	161	1	US-08-227-108-15	Sequence 49, Appl
	558	13	7.2	162	2	US-09-073-674-15	Sequence 49, Appl
	559	13	7.2	165	2	US-09-332-319-22	Sequence 277, App
	560	13	7.2	165	2	US-08-484-956-49	Sequence 55, App
	561	13	7.2	173	4	US-08-757-653-49	Sequence 442, App
	562	13	7.2	175	4	US-08-520-946-49	Sequence 7, Appl
	563	13	7.2	183	4	US-09-025-769B-277	Sequence 8, Appl
	564	13	7.2	187	5	PCT-US93-06251-55	Sequence 8, Appl
	565	13	7.2	198	1	US-09-060-756-442	Sequence 8, Appl
	566	13	7.2	198	1	US-08-734-925-7	Sequence 9, Appl
	567	13	7.2	198	1	US-08-276-852-8	Sequence 8, Appl
	568	13	7.2	198	1	US-08-276-852-9	Sequence 8, Appl
	569	13	7.2	198	1	US-08-899-575-8	Sequence 8, Appl
c	570	13	7.2	198	1	US-08-899-575-9	Sequence 9, Appl
	571	13	7.2	198	1	US-08-899-575-8	Sequence 8, Appli
c	572	13	7.2	198	1	US-08-899-575-9	Sequence 9, Appli
	573	13	7.2	198	1	PCT-US95-08743-8	Sequence 8, Appli
	574	13	7.2	198	5	PCT-US95-08743-9	Sequence 9, Appli
c	575	13	7.2	200	1	US-07-623-953-6	Sequence 6, Appli
c	576	13	7.2	201	1	US-08-276-852-41	Sequence 41, Appl
c	577	13	7.2	201	1	US-08-133-011-114	Sequence 114, App
	578	13	7.2	201	1	US-08-322-730A-114	Sequence 114, App
	579	13	7.2	201	1	US-08-387-874-87	Sequence 87, App
	580	13	7.2	201	1	US-08-899-575-41	Sequence 41, Appl
	581	13	7.2	201	1	US-08-899-575-41	Sequence 114, App
	582	13	7.2	201	1	US-08-383-619-114	Sequence 114, App
	583	13	7.2	201	1	US-08-907-739-114	Sequence 87, Appl
	584	13	7.2	201	4	US-09-729-597-114	Sequence 41, Appl
	585	13	7.2	201	5	PCT-US93-08364-87	Sequence 114, App
	586	13	7.2	201	5	PCT-US95-08743-41	Sequence 114, App
	587	13	7.2	203	2	PCT-US93-06251-49	Sequence 87, Appl
	588	13	7.2	206	2	US-08-254-359A-32	Sequence 41, Appl
	589	13	7.2	206	3	US-08-471-066B-32	Sequence 32, Appl
c	590	13	7.2	206	2	US-08-484-956-50	Sequence 32, Appl
c	591	13	7.2	206	2	US-08-757-653-32	Sequence 50, Appl
c	592	13	7.2	206	2	US-08-757-653-50	Sequence 32, Appl
c	593	13	7.2	206	2	US-08-599-491-32	Sequence 32, Appl
c	594	13	7.2	206	2	US-08-756-386-32	Sequence 50, Appl
c	595	13	7.2	206	3	US-08-823-516-27	Sequence 32, Appl
c	596	13	7.2	206	3	US-08-682-853A-32	Sequence 32, Appl
c	597	13	7.2	206	3	US-08-759-038-32	Sequence 32, Appl
c	598	13	7.2	206	3	US-08-758-314-32	Sequence 32, Appl
c	599	13	7.2	206	3	US-08-350-309-32	Sequence 32, Appl
c	600	13	7.2	206	3	US-08-520-946-32	Sequence 32, App
c	601	13	7.2	206	3	US-08-520-946-50	Sequence 231, App
c	602	13	7.2	208	3	US-09-404-879A-230	Sequence 520, App
c	603	13	7.2	208	3	US-09-404-879A-231	Sequence 603, App
c	604	13	7.2	215	4	US-09-060-756-520	Sequence 11, Appl
	605	13	7.2	217	4	US-09-060-756-603	Sequence 11, Appl
c	606	13	7.2	217	1	US-08-276-852-11	Sequence 11, Appl
c	607	13	7.2	220	1	US-08-276-852-12	Sequence 12, Appl
c	608	13	7.2	220	1	US-08-899-575-11	Sequence 11, Appl
c	609	13	7.2	217	1	US-08-276-852-11	Sequence 11, Appl
c	610	13	7.2	220	1	US-08-276-852-12	Sequence 12, Appl
c	611	13	7.2	220	1	US-08-899-575-11	Sequence 11, Appl

	SEQ			Pos	App	Patent No.	Reference
c	612	13	7.2	220	1	US-08-899-575-12	Sequence 12, Appl
	613	13	7.2	220	1	US-08-899-575-11	Sequence 11, Appl
c	614	13	7.2	220	1	US-08-899-575-12	Sequence 12, Appl
	615	13	7.2	220	5	PCT-US95-08743-11	Sequence 11, Appl
	616	13	7.2	220	5	PCT-US95-08743-12	Sequence 12, Appl
c	617	13	7.2	221	4	US-08-531-132-1	Sequence 1, Appl
	618	13	7.2	233	4	US-09-332-319-30	Sequence 100, App
	619	13	7.2	236	4	US-09-397-787-100	Sequence 1, Appli
	620	13	7.2	246	1	US-08-165-304A-1	Sequence 1, Appli
c	621	13	7.2	249	1	US-08-439-132-1	Sequence 10, App
	622	13	7.2	249	6	US-09-070-842A-10	Patent No. 5223407
	623	13	7.2	249	1	5223407-1	Sequence 6, Appl
	624	13	7.2	250	1	US-08-505-691-6	Sequence 28, Appl
	625	13	7.2	250	5	US-08-809-267-28	Sequence 19, Appl
	626	13	7.2	250	4	PCT-US95-13662A-28	Sequence 19, Appl
	627	13	7.2	261	4	US-08-991-789A-19	Sequence 19, Appl
	628	13	7.2	261	4	US-09-062-451-19	Sequence 1, Appli
	629	13	7.2	261	1	US-09-598-326-19	Sequence 20, Appl
c	630	13	7.2	264	4	US-08-182-961B-1	Sequence 12, Appl
c	631	13	7.2	264	1	US-09-007-678B-1	Sequence 47, Appl
	632	13	7.2	267	4	US-09-332-319-20	Sequence 1, Appli
	633	13	7.2	270	1	US-08-181-629A-12	Sequence 4, Appl
	634	13	7.2	294	4	US-09-254-352B-47	Sequence 8, Appl
	635	13	7.2	295	3	US-09-121-979-1	Sequence 8, Appli
	636	13	7.2	295	4	US-09-332-319-1	Sequence 8, Appli
	637	13	7.2	296	4	US-08-715-941-4	Sequence 11, Appl
c	638	13	7.2	297	1	US-07-602-824A-8	Sequence 45, Appl
	639	13	7.2	297	1	US-07-983-451-8	Sequence 63, Appl
	640	13	7.2	297	1	US-08-261-577-11	Sequence 21, Appl
	641	13	7.2	301	4	US-09-254-352B-45	Sequence 49, Appl
	642	13	7.2	304	4	US-09-712-016-63	Sequence 23, Appl
	643	13	7.2	305	3	US-09-299-843A-21	Sequence 23, Appl
	644	13	7.2	305	4	US-09-088-337B-21	Sequence 25, Appl
c	645	13	7.2	305	4	US-09-254-352B-49	Sequence 49, Appl
	646	13	7.2	321	1	US-07-920-519-23	Sequence 23, Appl
	647	13	7.2	321	3	US-08-314-586-23	Sequence 23, Appl
	648	13	7.2	321	3	US-08-115-753-25	Sequence 25, Appl
c	649	13	7.2	326	4	US-09-171-461-39	Sequence 39, Appl
	650	13	7.2	326	4	US-09-058-746-7	Sequence 7, Appl
	651	13	7.2	326	4	US-09-438-142-7	Sequence 7, Appli
	652	13	7.2	327	5	PCT-US93-02612-8	Sequence 8, Appli
	653	13	7.2	341	4	US-09-060-756-355	Sequence 355, App
c	654	13	7.2	354	4	US-09-385-982-498	Sequence 498, App
	655	13	7.2	357	4	US-09-525-046-3	Sequence 3, Appli
c	656	13	7.2	360	2	US-08-182-173A-1	Sequence 1, Appli
c	657	13	7.2	362	2	US-08-875-972-16	Sequence 16, Appl
	658	13	7.2	371	4	US-09-404-879A-365	Sequence 365, App
c	659	13	7.2	372	3	US-09-070-842A-7	Sequence 7, Appl
	660	13	7.2	375	4	US-09-070-842A-11	Sequence 11, Appl
	661	13	7.2	376	2	US-09-712-016-61	Sequence 61, Appl
	662	13	7.2	380	3	US-08-735-609-5	Sequence 5, Appl
	663	13	7.2	380	3	US-08-735-609-5	Sequence 5, Appli
	664	13	7.2	380	2	US-08-315-372-5	Sequence 5, Appl
	665	13	7.2	380	3	US-09-244-752-5	Sequence 5, Appl
	666	13	7.2	380	4	US-09-245-497-5	Sequence 5, Appl
	667	13	7.2	387	4	US-09-562-919-5	Sequence 481, App
	668	13	7.2	397	4	US-09-060-756-481	Sequence 250, App
c	669	13	7.2	414	4	US-09-199-637A-250	Sequence 9, Appl
	670	13	7.2	418	4	US-08-576-202-9	Sequence 9, Appl
	671	13	7.2	418	5	PCT-US95-16766-9	Sequence 230, App
	672	13	7.2	425	4	US-09-328-111-230	Sequence 539, App
	673	13	7.2	431	4	US-09-060-756-539	Sequence 7, Appli
	674	13	7.2	439	4	US-09-218-467B-7	Sequence 373, App
	675	13	7.2	443	4	US-09-060-756-373	Sequence 149, App
	676	13	7.2	445	4	US-09-397-787-149	Sequence 70, App
	677	13	7.2	447	4	US-09-060-756-70	Sequence 553, App
	678	13	7.2	453	4	US-09-060-756-553	Sequence 371, App
	679	13	7.2	455	4	US-09-060-756-371	Sequence 591, App
	680	13	7.2	457	4	US-09-060-756-591	Sequence 311, App
	681	13	7.2	458	4	US-09-060-756-311	Sequence 551, App
	682	13	7.2	458	4	US-09-060-756-551	Sequence 68, App
	683	13	7.2	465	4	US-09-060-756-68	Sequence 555, App
	684	13	7.2	466	4	US-09-060-756-555	Sequence 555, App

us-09-550-163-1_copy_80_260.oligo.rni

Left column

	Index	Score	Len	Line	Patent Application No.	N	Sequence
	685	7.2	13	469	US-08-990-823-65	4	Sequence 65, Appl
	686	7.2	13	473	US-09-060-756-288	4	Sequence 288, App
	687	7.2	13	473	US-09-060-756-333	4	Sequence 333, App
c	688	7.2	13	479	US-09-254-352B-39	4	Sequence 39, Appl
	689	7.2	13	480	US-08-288-663A-14	2	Sequence 14, Appl
	690	7.2	13	481	US-08-280-116-23	4	Sequence 23, Appl
	691	7.2	13	488	US-09-174-943-3	4	Sequence 3, Appli
	692	7.2	13	490	US-08-896-164-66	4	Sequence 66, Appl
	693	7.2	13	504	US-08-768-550-12	4	Sequence 12, Appl
c	694	7.2	13	505	US-08-768-550-10	2	Sequence 10, Appl
c	695	7.2	13	506	US-08-768-550-11	2	Sequence 11, Appl
c	696	7.2	13	519	US-08-068-140A-3	2	Sequence 3, Appli
c	697	7.2	13	535	US-08-463-115-57	1	Sequence 57, Appl
c	698	7.2	13	535	US-08-465-388-57	1	Sequence 1, Appli
c	699	7.2	13	536	US-08-329-704-1	1	Sequence 1, Appli
c	700	7.2	13	536	US-08-486-117-1	2	Sequence 1, Appli
c	701	7.2	13	536	US-08-477-537-1	2	Sequence 1, Appli
c	702	7.2	13	545	US-09-227-347-125	1	Sequence 125, App
c	703	7.2	13	551	US-08-165-304A-3	4	Sequence 3, Appli
	704	7.2	13	563	US-08-463-115-73	1	Sequence 73, Appl
c	705	7.2	13	563	US-08-465-388-73	1	Sequence 56, Appl
c	706	7.2	13	565	US-09-712-016-56	1	Sequence 1049, Ap
c	707	7.2	13	570	US-09-134-001C-1049	4	Sequence 551, App
	708	7.2	13	573	US-09-328-111-551	4	Sequence 2508, Ap
	709	7.2	13	576	US-09-134-001C-2508	4	Sequence 773, App
c	710	7.2	13	580	US-09-328-111-773	4	Sequence 92, Appl
	711	7.2	13	585	US-09-123-912-92	4	Sequence 92, Appl
c	712	7.2	13	585	US-09-643-597-92	4	Sequence 1, Appli
	713	7.2	13	588	US-08-965-904B-1	3	Sequence 107, Appl
	714	7.2	13	593	US-08-961-083-107	3	Sequence 154, App
c	715	7.2	13	596	US-09-222-575-154	4	Sequence 23, Appl
c	716	7.2	13	599	US-09-078-294-23	4	Patent No. 5290690
c	717	7.2	13	600	5290690-4	6	Sequence 68, Appl
c	718	7.2	13	611	US-08-896-164-68	4	Sequence 1, Appli
	719	7.2	13	614	US-08-217-327-1	1	Sequence 4, Appl
c	720	7.2	13	614	US-08-241-943-4	1	Sequence 4, Appli
c	721	7.2	13	614	US-08-378-588-4	1	Sequence 4, Appli
c	722	7.2	13	614	US-08-811-094-4	1	Sequence 4, Appli
c	723	7.2	13	615	PCT-US94-11121-4	5	Sequence 30, Appl
c	724	7.2	13	617	US-09-385-982-30	4	Sequence 522, App
	725	7.2	13	621	US-09-385-982-522	4	Sequence 26, Appl
c	726	7.2	13	621	US-08-989-251-26	3	Sequence 26, Appl
	727	7.2	13	621	US-09-340-250-26	3	Sequence 26, Appl
	728	7.2	13	623	US-09-528-108-26	3	Sequence 167, Appl
c	729	7.2	13	625	US-09-385-982-167	4	Sequence 24, Appl
c	730	7.2	13	627	US-08-867-352-24	3	Sequence 2, Appli
c	731	7.2	13	627	US-08-981-030-2	4	Sequence 12, Appl
c	732	7.2	13	627	US-08-981-030-12	4	Sequence 13, Appl
c	733	7.2	13	641	US-08-981-030-13	4	Sequence 247, Appl
	734	7.2	13	650	US-08-936-165A-247	4	Sequence 729, App
c	735	7.2	13	659	US-09-060-756-729	4	Sequence 1118, Ap
c	736	7.2	13	660	US-08-998-416-1118	4	Sequence 1, Appli
c	737	7.2	13	660	US-09-498-270-1	1	Sequence 4, Appli
c	738	7.2	13	663	US-08-981-030-4	4	Sequence 2236, Ap
	739	7.2	13	681	US-09-134-001C-2236	4	Sequence 2649, Ap
c	740	7.2	13	682	US-09-134-001C-2649	2	Sequence 23, Appl
c	741	7.2	13	687	US-08-706-702-23	3	Sequence 23, Appl
	742	7.2	13	687	US-08-706-706-23	4	Sequence 1084, App
c	743	7.2	13	691	US-08-998-416-1084	2	Sequence 24, Appl
c	744	7.2	13	691	US-08-706-702-24	4	Sequence 24, Appl
c	745	7.2	13	691	US-08-706-706-24	5	Sequence 12, Appl
c	746	7.2	13	693	PCT-US95-05423-12	6	Patent No. 5175255
	747	7.2	13	701	5175255-3	4	Sequence 136, Appl
c	748	7.2	13	701	US-08-998-416-136	3	Sequence 11, Appl
c	749	7.2	13	703	US-08-862-540-11	3	Sequence 12, Appl
c	750	7.2	13	705	US-08-862-540-12	3	Sequence 11, Appl
c	751	7.2	13	705	US-09-633-043-11	3	Sequence 26, Appl
c	752	7.2	13	705	US-08-706-702-26	3	Sequence 26, Appl
c	753	7.2	13	705	US-08-706-706-26	3	Sequence 26, Appl

Right column

	Index	Score	Len	Line	Patent Application No.	N	Sequence
	758	7.2	13	715	US-08-545-809A-11	3	Sequence 11, Appl
c	759	7.2	13	716	US-08-998-416-55	3	Sequence 55, Appl
c	760	7.2	13	717	US-08-706-702-25	2	Sequence 25, Appl
c	761	7.2	13	717	US-08-706-706-25	3	Sequence 25, Appl
c	762	7.2	13	729	US-09-712-016-58	4	Sequence 58, Appl
c	763	7.2	13	730	US-09-712-016-59	4	Sequence 59, Appl
	764	7.2	13	731	5194596-14	6	Patent No. 5194596
	765	7.2	13	736	US-08-862-540-8	3	Sequence 8, Appli
	766	7.2	13	736	US-08-862-540-10	3	Sequence 10, Appl
c	767	7.2	13	736	US-09-633-043-8	4	Sequence 8, Appli
	768	7.2	13	738	US-09-633-043-10	4	Sequence 10, Appl
	769	7.2	13	738	US-08-726-462B-3	1	Sequence 3, Appli
	770	7.2	13	738	US-09-046-203-3	2	Sequence 3, Appli
c	771	7.2	13	739	US-09-272-104-3	4	Sequence 3, Appli
c	772	7.2	13	741	US-09-272-097-3	4	Sequence 3, Appli
c	773	7.2	13	751	5219739-14	6	Patent No. 5219739
	774	7.2	13	752	US-09-370-838-103	4	Sequence 103, App
	775	7.2	13	762	US-09-019-931-1	4	Sequence 1, Appli
c	776	7.2	13	769	US-08-976-259-108	4	Sequence 108, App
c	777	7.2	13	773	US-09-134-001C-2031	4	Sequence 2031, Ap
c	778	7.2	13	774	US-08-858-207A-154	3	Sequence 154, Ap
c	779	7.2	13	774	US-08-934-131-2	3	Sequence 2, Appli
c	780	7.2	13	782	US-08-765-340-1	3	Sequence 1, Appli
	781	7.2	13	783	US-08-896-164-56	4	Sequence 56, Appl
c	782	7.2	13	784	US-09-020-956-30	4	Sequence 30, Appl
	783	7.2	13	785	US-09-030-607-30	4	Sequence 30, Appl
	784	7.2	13	786	US-09-605-785-30	4	Sequence 30, Appl
	785	7.2	13	787	US-09-439-313-30	4	Sequence 30, Appl
	786	7.2	13	787	US-09-352-616A-30	4	Sequence 30, Appl
	787	7.2	13	787	US-09-232-149A-30	4	Sequence 30, Appl
c	788	7.2	13	793	US-09-020-956-33	4	Sequence 33, Appl
	789	7.2	13	793	US-09-030-607-33	4	Sequence 33, Appl
	790	7.2	13	793	US-09-605-785-33	4	Sequence 33, Appl
	791	7.2	13	793	US-09-439-313-33	4	Sequence 33, Appl
c	792	7.2	13	793	US-09-352-616A-33	4	Sequence 33, Appl
c	793	7.2	13	794	US-09-232-149A-33	4	Sequence 33, Appl
c	794	7.2	13	795	US-08-887-534A-75	4	Sequence 75, Appl
c	795	7.2	13	796	US-08-887-534A-77	4	Sequence 77, Appl
	796	7.2	13	799	US-09-975-316-50	2	Sequence 50, Appl
	797	7.2	13	799	US-09-007-119-9	4	Sequence 9, Appli
c	798	7.2	13	801	US-09-615-192A-50	2	Sequence 50, Appl
	799	7.2	13	801	US-09-767-026-12	4	Sequence 12, Appl
	800	7.2	13	804	US-08-961-527-372	2	Sequence 372, App
	801	7.2	13	807	5206352-1	6	Patent No. 5206352
c	802	7.2	13	807	US-08-961-527-337	4	Sequence 337, App
c	803	7.2	13	807	US-09-020-956-1	4	Sequence 1, Appli
c	804	7.2	13	813	US-09-030-607-1	4	Sequence 1, Appli
	805	7.2	13	814	US-09-605-785-1	4	Sequence 1, Appli
c	806	7.2	13	814	US-09-439-313-1	4	Sequence 1, Appli
c	807	7.2	13	814	US-09-352-616A-1	4	Sequence 1, Appli
c	808	7.2	13	814	US-09-232-149A-1	4	Sequence 1, Appli
c	809	7.2	13	816	US-09-020-956-2	4	Sequence 2, Appli
c	810	7.2	13	816	US-09-030-607-2	4	Sequence 2, Appli
c	811	7.2	13	816	US-09-605-785-2	4	Sequence 2, Appli
c	812	7.2	13	816	US-09-439-313-2	4	Sequence 2, Appli
c	813	7.2	13	816	US-09-352-616A-2	4	Sequence 2, Appli
c	814	7.2	13	827	US-09-221-017B-265	4	Sequence 265, App
	815	7.2	13	828	US-09-030-956-4	4	Sequence 4, Appli
c	816	7.2	13	828	US-09-030-607-4	4	Sequence 4, Appli
	817	7.2	13	828	US-09-605-785-4	4	Sequence 4, Appli
	818	7.2	13	828	US-09-439-313-4	4	Sequence 4, Appli
c	819	7.2	13	837	US-09-352-616A-4	4	Sequence 4, Appli
	820	7.2	13	840	US-09-232-149A-4	4	Sequence 4, Appli
c	821	7.2	13	841	US-09-134-001C-1444	4	Sequence 1444, Ap
	822	7.2	13	846	US-09-134-001C-1385	4	Sequence 1385, Ap
	823	7.2	13	861	US-08-943-731-148	4	Sequence 148, App
	824	7.2	13	862	US-09-292-858B-9	4	Sequence 9, Appli
c	825	7.2	13	865	US-08-953-326-8	4	Sequence 8, Appli
	826	7.2	13	868	US-08-961-527-285	4	Sequence 285, App
	827	7.2	13		US-08-341-568-4	2	Sequence 4, Appli
c	828	7.2	13		US-08-911-020-4	2	Sequence 4, Appli
c	829	7.2	13		US-08-387-845-3	1	Sequence 3, Appli
c	830	7.2	13				

Left portion (entries 831–903):

Idx	c	n	ver	len	pos	Patent/Publication No.	Sequence
831		2	7.2	13	868	US-08-778-275-3	Sequence 3, Appli
832		3	7.2	13	868	US-08-867-152-3	Sequence 3, Appli
833	c	1	7.2	13	876	US-08-928-284-1	Sequence 1, Appl
834		4	7.2	13	904	US-09-171-209-40	Sequence 40, Appl
835	c	4	7.2	13	906	US-09-134-001C-2072	Sequence 2072, Ap
836		4	7.2	13	908	US-08-991-789A-46	Sequence 46, Appl
837		4	7.2	13	908	US-08-991-789A-46	Sequence 46, Appl
838		4	7.2	13	908	US-09-062-451-46	Sequence 46, Appl
839	c	4	7.2	13	908	US-09-062-451-46	Sequence 46, Appl
840		4	7.2	13	908	US-09-598-326-46	Sequence 46, Appl
841		4	7.2	13	908	US-09-598-326-46	Sequence 46, Appl
842		4	7.2	13	909	US-09-134-001C-334	Sequence 334, App
843		4	7.2	13	924	US-09-484-970B-42	Sequence 42, Appl
844		2	7.2	13	941	US-08-121-436A-3	Sequence 3, Appl
845		2	7.2	13	949	US-08-713-000-4	Sequence 4, Appli
846		2	7.2	13	949	US-08-975-316-4	Sequence 4, Appli
847		4	7.2	13	949	US-09-211-710-4	Sequence 4, Appli
848		4	7.2	13	950	US-09-615-192A-4	Sequence 4, Appli
849		3	7.2	13	955	US-08-984-550-1	Sequence 1, Appl
850		4	7.2	13	955	US-09-641-638-21	Sequence 21, Appl
851		4	7.2	13	981	US-09-641-638-22	Sequence 22, Appl
852		4	7.2	13	984	US-08-986-768-1	Sequence 1, Appli
853		4	7.2	13	984	US-08-986-768-3	Sequence 3, Appli
854	c	1	7.2	13	984	US-08-420-235B-34	Sequence 34, Appl
855		4	7.2	13	999	US-08-793-624-34	Sequence 34, Appl
856		5	7.2	13	999	PCT-US95-10194-34	Sequence 34, Appl
857		2	7.2	13	1001	US-08-860-882A-67	Sequence 67, Appl
858		4	7.2	13	1001	US-09-011-769A-50	Sequence 50, Appl
859	c	4	7.2	13	1011	US-09-641-638-304	Sequence 304, App
860		4	7.2	13	1026	US-09-641-638-354	Sequence 354, App
861		1	7.2	13	1053	US-08-713-254-1	Sequence 1, Appli
862		1	7.2	13	1053	US-08-855-767-1	Sequence 1, Appli
863		4	7.2	13	1059	US-09-134-001C-2398	Sequence 2398, Ap
864	c	2	7.2	13	1059	US-08-860-882A-64	Sequence 64, Appl
865		3	7.2	13	1059	US-09-011-769A-46	Sequence 46, Appl
866		4	7.2	13	1064	US-08-860-882A-74	Sequence 74, Appl
867		4	7.2	13	1073	US-08-860-882A-77	Sequence 77, Appl
868		4	7.2	13	1073	US-08-276-531-74	Sequence 74, Appl
869		4	7.2	13	1097	US-09-011-769A-59	Sequence 59, Appl
870	c	4	7.2	13	1097	US-09-011-769A-63	Sequence 63, Appl
871		1	7.2	13	1109	US-09-149-922-3	Sequence 3, Appli
872		1	7.2	13	1140	US-08-753-159A-2	Sequence 2, Appli
873		3	7.2	13	1140	US-09-133-735-2	Sequence 2, Appli
874		2	7.2	13	1164	US-08-977-306-7	Sequence 7, Appli
875		4	7.2	13	1180	US-09-149-476-140	Sequence 140, App
876		1	7.2	13	1180	US-08-704-548-1	Sequence 4, Appli
877	c	3	7.2	13	1189	US-09-023-173-4	Sequence 4, Appli
878		4	7.2	13	1189	US-09-173-270-1	Sequence 1, Appli
879		2	7.2	13	1194	US-09-134-001C-411	Sequence 411, App
880		3	7.2	13	1194	US-08-897-236-10	Sequence 10, Appl
881		3	7.2	13	1197	US-08-167-874-10	Sequence 10, Appl
882		1	7.2	13	1217	US-07-781-034-4	Sequence 4, Appl
883		1	7.2	13	1217	PCT-US92-0328-4	Sequence 4, Appli
884		5	7.2	13	1217	US-08-447-702-4	Sequence 4, Appli
885		1	7.2	13	1228	US-08-465-615-4	Sequence 4, Appli
886		2	7.2	13	1236	US-08-829-026A-5	Sequence 5, Appli
887		1	7.2	13	1237	US-08-726-462B-1	Sequence 1, Appli
888		2	7.2	13	1256	US-08-642-330-1	Sequence 1, Appli
889		1	7.2	13	1256	US-09-046-203-1	Sequence 1, Appli
890		2	7.2	13	1263	US-09-272-104-1	Sequence 1, Appli
891		3	7.2	13	1263	US-09-272-097-1	Sequence 3, Appli
892		2	7.2	13	1269	US-08-288-663A-3	Sequence 9, Appli
893		2	7.2	13	1269	US-08-423-439-9	Sequence 2, Appli
894		1	7.2	13	1269	US-08-798-000-2	Sequence 7, Appli
895		2	7.2	13	1272	US-09-553-498-7	Sequence 56, Appl
896		2	7.2	13		US-09-618-869-7	Sequence 38, Appl
897		2	7.2	13		US-08-860-882A-56	Sequence 1, Appli
898	c	1	7.2	13		US-09-011-769A-38	Sequence 1, Appli
899		1	7.2	13		US-08-621-700-1	Sequence 5, Appli
900		1	7.2	13		US-08-929-940-1	Sequence 1, Appli
901		5	7.2	13		PCT-US95-03940-1	Sequence 2, Appli
902		1	7.2	13		US-08-908-758-3	Sequence 3, Appli
903	c	4	7.2	13			

Right portion (entries 904–976):

Idx	c	len	ver	pos	n	Patent/Publication No.	Sequence
904	c	13	7.2	1284	2	US-08-860-882A-71	Sequence 71, Appl
905	c	13	7.2	1284	4	US-09-011-769A-55	Sequence 55, Appl
906	c	13	7.2	1293	4	US-08-313-412-2	Sequence 2, Appli
907		13	7.2	1304	4	US-09-372-422A-15	Sequence 15, Appl
908		13	7.2	1320	3	US-09-989-251-28	Sequence 28, Appl
909		13	7.2	1320	4	US-09-340-250-28	Sequence 28, Appl
910		13	7.2	1320	4	US-09-528-108-28	Sequence 28, Appl
911	c	13	7.2	1333	1	US-09-496-632C-2	Sequence 2, Appli
912		13	7.2	1353	3	US-08-214-580A-1	Sequence 1, Appli
913	c	13	7.2	1355	4	US-08-974-022-5	Sequence 5, Appli
914	c	13	7.2	1355	4	US-08-795-445A-5	Sequence 5, Appli
915	c	13	7.2	1355	4	US-08-795-447A-5	Sequence 5, Appli
916		13	7.2	1355	4	US-08-974-186-5	Sequence 5, Appli
917		13	7.2	1355	4	US-08-795-446B-5	Sequence 5, Appli
918	c	13	7.2	1362	6	US-08-706-945D-127	Sequence 127, App
919		13	7.2	1379	4	5290690-8	Patent No. 5290690
920		13	7.2	1379	4	US-09-553-498-5	Sequence 5, Appli
921		13	7.2	1380	4	US-09-618-869-5	Sequence 5, Appli
922	c	13	7.2	1380	4	US-09-305-984-19	Sequence 19, Appl
923		13	7.2	1410	2	US-09-073-541A-19	Sequence 19, Appl
924		13	7.2	1412	4	US-09-378-088A-81	Sequence 81, Appl
925		13	7.2	1422	1	US-08-378-617A-7	Sequence 7, Appli
926	c	13	7.2	1423	5	US-09-485-077A-15	Sequence 15, Appl
927		13	7.2	1423	2	US-08-214-580A-2	Sequence 2, Appli
928		13	7.2	1440	4	PCT-US95-07554-5	Sequence 5, Appli
929		13	7.2	1440	4	US-08-743-637B-174	Sequence 174, App
930		13	7.2	1478	4	US-08-526-840B-174	Sequence 174, App
931		13	7.2	1491	1	US-09-149-922-6	Sequence 6, Appli
932		13	7.2	1491	4	US-08-740-223A-21	Sequence 21, Appl
933		13	7.2	1493	1	US-09-709-188-21	Sequence 21, Appl
934		13	7.2	1493	6	US-08-340-820-24	Sequence 24, Appl
935		13	7.2	1493	4	US-08-593-535-24	Sequence 24, Appl
936	c	13	7.2	1496	4	5340034-5	Patent No. 5340934
937		13	7.2	1496	1	US-08-868-699A-3	Sequence 3, Appli
938		13	7.2	1500	1	US-09-757-014-3	Sequence 3, Appli
939		13	7.2	1500	2	US-08-740-223A-23	Sequence 23, Appl
940		13	7.2	1518	4	US-09-709-188-23	Sequence 23, Appl
941		13	7.2	1518	4	US-08-114-695A-6	Sequence 6, Appli
942	c	13	7.2	1518	1	US-08-815-667-9	Sequence 9, Appli
943		13	7.2	1522	1	US-08-815-667-13	Sequence 13, Appl
944		13	7.2	1527	2	US-08-453-702B-225	Sequence 225, App
945		13	7.2	1527	4	US-07-956-483-26	Sequence 26, Appl
946		13	7.2	1542	4	US-08-472-240A-19	Sequence 19, Appl
947		13	7.2	1543	4	US-09-276-531-83	Sequence 83, Appl
948		13	7.2	1550	4	US-09-227-357-26	Sequence 26, Appl
949	c	13	7.2	1553	4	US-08-943-731-70	Sequence 70, Appl
950		13	7.2	1558	4	US-09-364-230-31	Sequence 31, Appl
951		13	7.2	1558	4	US-08-343-101A-4	Sequence 4, Appli
952		13	7.2	1561	3	US-09-183-688-4	Sequence 4, Appli
953	c	13	7.2	1572	3	US-08-965-850-1	Sequence 1, Appli
954		13	7.2	1572	1	US-08-041-219A-5	Sequence 5, Appli
955		13	7.2	1578	1	US-08-417-122-5	Sequence 5, Appli
956		13	7.2	1578	1	US-07-885-970A-23	Sequence 23, Appl
957		13	7.2	1578	3	US-08-298-687A-23	Sequence 23, Appl
958		13	7.2	1584	3	US-08-298-829-23	Sequence 23, Appl
959		13	7.2	1584	4	US-08-747-221B-51	Sequence 51, Appl
960		13	7.2	1584	4	US-08-747-221B-52	Sequence 51, Appl
961	c	13	7.2	1597	3	US-09-005-051-51	Sequence 52, Appl
962		13	7.2	1599	4	US-09-005-051-52	Sequence 52, Appl
963		13	7.2	1599	4	US-09-038-832-3	Sequence 3, Appli
964	c	13	7.2	1611	3	US-09-256-465-1	Sequence 1, Appli
965		13	7.2	1618	1	US-09-167-322-3	Sequence 3, Appli
966		13	7.2	1618	1	US-08-061-062A-5	Sequence 5, Appli
967		13	7.2	1618	1	US-08-536-150-5	Sequence 5, Appli
968		13	7.2	1619	1	US-07-885-970A-22	Sequence 22, Appl
969	c	13	7.2	1631	2	US-08-298-687A-22	Sequence 22, Appl
970		13	7.2	1631	2	US-08-298-829-22	Sequence 22, Appl
971		13	7.2	1631	2	US-08-507-455-1	Sequence 1, Appli
972		13	7.2	1631	2	US-08-583-318-4	Sequence 4, Appli
973	c	13	7.2	1647	1	US-08-498-263-5	Sequence 5, Appli
974		13	7.2			US-08-977-060-5	Sequence 5, Appli
975	c	13	7.2			US-09-014-583-2	Sequence 2, Appli
976	c	13	7.2			US-07-885-970A-20	Sequence 20, Appl

```
c 977   13  7.2  1647  1  US-08-298-687A-20     Sequence 20, Appl
c 978   13  7.2  1647  1  US-08-298-829-20      Sequence 20, Appl
c 979   13  7.2  1650  1  US-09-399-561-5       Sequence 5, Appli
c 980   13  7.2  1666  4  US-09-149-922-4       Sequence 4, Appli
c 981   13  7.2  1668  1  US-07-885-970A-21     Sequence 21, Appl
  982   13  7.2  1668  1  US-08-298-687A-21     Sequence 21, Appl
  983   13  7.2  1668  1  US-08-298-829-21      Sequence 21, Appl
  984   13  7.2  1672  1  US-08-530-797-13      Sequence 13, Appl
  985   13  7.2  1672  1  US-08-787-335-13      Sequence 13, Appl
  986   13  7.2  1675  4  US-09-027-900-7       Sequence 7, Appli
  987   13  7.2  1683  4  US-09-385-742B-5      Sequence 5, Appli
  988   13  7.2  1702  1  US-08-261-832A-14     Sequence 14, Appl
  989   13  7.2  1702  5  PCT-US95-07744A-14    Sequence 14, Appl
  990   13  7.2  1707  4  US-09-526-993-5       Sequence 5, Appli
  991   13  7.2  1713  1  US-09-000-062-1       Sequence 1, Appli
c 992   13  7.2  1728  2  US-08-532-795-24      Sequence 24, Appl
  993   13  7.2  1738  1  US-07-971-160-9       Sequence 9, Appli
  994   13  7.2  1738  1  US-08-336-241-9       Sequence 9, Appli
  995   13  7.2  1738  2  US-08-465-273-9       Sequence 9, Appli
  996   13  7.2  1738  2  US-09-119-024-9       Sequence 9, Appli
  997   13  7.2  1738  2  US-08-417-226-9       Sequence 9, Appli
  998   13  7.2  1738  4  US-09-196-131-9       Sequence 9, Appli
  999   13  7.2  1738  4  US-08-643-732-9       Sequence 9, Appli
c1000   13  7.2  1749  2  US-08-899-244-1       Sequence 1, Appli
```

ALIGNMENTS

RESULT 1
US-08-570-311-5/c
; Sequence 5, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
. APPLICANT: Progulske-Fox, Ann
.. APPLICANT: Tumwasorn, Somying
.. APPLICANT: Lepine, Guylaine
.. APPLICANT: Han, Naiming
.. APPLICANT: Lantz, Marilyn
.. APPLICANT: Patti, Joseph
.. TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
.. TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
.. NUMBER OF SEQUENCES: 29
.. CORRESPONDENCE ADDRESS:
... ADDRESSEE: Ted W. Whitlock
... STREET: 2421 N.W. 41st Street, Suite A-1
... CITY: Gainesville
... STATE: FL
... COUNTRY: USA
... ZIP: 32606
.. COMPUTER READABLE FORM:
... MEDIUM TYPE: Floppy disk
... COMPUTER: IBM PC compatible
... OPERATING SYSTEM: PC-DOS/MS-DOS
... SOFTWARE: PatentIn Release #1.0, Version #1.25
.. CURRENT APPLICATION DATA:
... APPLICATION NUMBER: US/08/570,311
... FILING DATE:
... CLASSIFICATION: 424
.. PRIOR APPLICATION DATA:
... APPLICATION NUMBER: US 08/353,485
... FILING DATE: 09-DEC-1994
... CLASSIFICATION: 424
.. PRIOR APPLICATION DATA:
... APPLICATION NUMBER: US 07/647,119
... FILING DATE: 25-JAN-1991
... CLASSIFICATION: 424
.. PRIOR APPLICATION DATA:
... APPLICATION NUMBER: US 07/241,640
... FILING DATE: 08-SEP-1988
.. ATTORNEY/AGENT INFORMATION:
... NAME: Whitlock, Ted W.
... REGISTRATION NUMBER: 36,965
... REFERENCE/DOCKET NUMBER: UF15.C3
.. TELECOMMUNICATION INFORMATION:
... TELEPHONE: (904) 375-8100
... TELEFAX: (904) 372-5800
.. INFORMATION FOR SEQ ID NO: 5:
... SEQUENCE CHARACTERISTICS:
.... LENGTH: 1841 base pairs
.... TYPE: nucleic acid
.... STRANDEDNESS: single
.... TOPOLOGY: linear
... MOLECULE TYPE: DNA (genomic)
... FEATURE:
.... NAME/KEY: CDS
.... LOCATION: 374..1424
US-08-570-311-5

```
Query Match            9.9%;  Score 18;  DB 1;  Length 1841;
Best Local Similarity 100.0%;  Pred. No. 1.7;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    24 GCTGGAAGACGTCTTCCG 41
         ||||||||||||||||||
Db  1627 GCTGGAAGACGTCTTCCG 1610
```

RESULT 2
US-08-353-485-5/c
; Sequence 5, Application US/08353485
; Patent No. 5830710
; GENERAL INFORMATION:
. APPLICANT: Progulske-Fox, Ann
.. APPLICANT: Tumwasorn, Somying
.. APPLICANT: Lepine, Guylaine
.. APPLICANT: Han, Naiming
.. APPLICANT: Lantz, Marilyn
.. APPLICANT: Patti, Joseph
.. TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
.. TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
.. NUMBER OF SEQUENCES: 12
.. CORRESPONDENCE ADDRESS:
... ADDRESSEE: Ted W. Whitlock
... STREET: 2421 N.W. 41st Street, Suite A-1
... CITY: Gainesville
... STATE: FL
... COUNTRY: USA
... ZIP: 32606
.. COMPUTER READABLE FORM:
... MEDIUM TYPE: Floppy disk
... COMPUTER: IBM PC compatible
... OPERATING SYSTEM: PC-DOS/MS-DOS
... SOFTWARE: PatentIn Release #1.0, Version #1.25
.. CURRENT APPLICATION DATA:
... APPLICATION NUMBER: US/08/353,485
... FILING DATE: 09-DEC-1994
... CLASSIFICATION: 424
.. PRIOR APPLICATION DATA:
... APPLICATION NUMBER: US 07/647,119
... FILING DATE: 25-JAN-1991
... CLASSIFICATION: 424
.. PRIOR APPLICATION DATA:
... APPLICATION NUMBER: US 07/241,640
... FILING DATE: 08-SEP-1988
.. ATTORNEY/AGENT INFORMATION:
... NAME: Whitlock, Ted W.
... REGISTRATION NUMBER: 36,965
... REFERENCE/DOCKET NUMBER: UF15.C2
.. TELECOMMUNICATION INFORMATION:
... TELEPHONE: (904) 375-8100
... TELEFAX: (904) 372-5800
.. INFORMATION FOR SEQ ID NO: 5:
... SEQUENCE CHARACTERISTICS:
.... LENGTH: 1841 base pairs
.... TYPE: nucleic acid

```
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
;      FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 374..1424
US-08-353-485-5

Query Match        9.9%;    Score 18;   DB 2;   Length 1841;
Best Local Similarity  100.0%;   Pred. No. 1.7;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       24 GCTGGAAGACGTCTTCCG 41
            ||||||||||||||||||
Db     1627 GCTGGAAGACGTCTTCCG 1610


RESULT 3
US-08-446-794A-3/c
; Sequence 3, Application US/08446794A
; Patent No. 5747327
; GENERAL INFORMATION:
;   APPLICANT: UEKI, JUN
;   APPLICANT: MORIOKA, SHINJI
;   TITLE OF INVENTION: PHOSPHOLIPASE D GENE ORIGINATED FROM
;   TITLE OF INVENTION: PLANT
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
;     STREET: P.O. BOX 747
;     CITY: FALLS CHURCH
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22040-0747
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/446,794A
;     FILING DATE:
;     CLASSIFICATION: 536
;   ATTORNEY/AGENT INFORMATION:
;     NAME: MURPHY JR, GERALD M
;     REGISTRATION NUMBER: 28,977
;     REFERENCE/DOCKET NUMBER: 0760-0203P
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-205-8000
;     TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2804 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 107..2542
US-08-446-794A-3

Query Match        9.9%;    Score 18;   DB 1;   Length 2804;
Best Local Similarity  100.0%;   Pred. No. 1.7;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       23 CGCTGGAAGACGTCTTCC 40
            ||||||||||||||||||
Db     2310 CGCTGGAAGACGTCTTCC 2293


RESULT 4
```

```
US-08-378-698-2/c
; Sequence 2, Application US/0378698
; Patent No. 5766914
; GENERAL INFORMATION:
;   APPLICANT: Deits, Thomas L.
;   TITLE OF INVENTION: METHOD OF PRODUCING AND PURIFYING ENZYMES
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Quarles & Brady
;     STREET: 411 East Wisconsin Avenue
;     CITY: Milwaukee
;     STATE: Wisconsin
;     COUNTRY: U.S.A.
;     ZIP: 53202-4497
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/378,698
;     FILING DATE:
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Baker, Jean C.
;     REGISTRATION NUMBER: 35,433
;     REFERENCE/DOCKET NUMBER: 660336.90489
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (414) 277-5709
;     TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 249 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-378-698-2

Query Match        9.4%;    Score 17;   DB 1;   Length 249;
Best Local Similarity  100.0%;   Pred. No. 5.6;
Matches  17;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       61 ATGGACAATTGGCGCCA 77
            |||||||||||||||||
Db       63 ATGGACAATTGGCGCCA 47


RESULT 5
PCT-US96-00728-2/c
; Sequence 2, Application PC/TUS9600728
; GENERAL INFORMATION:
;   APPLICANT: Deits, Thomas L.
;   TITLE OF INVENTION: METHOD OF PRODUCING AND PURIFYING ENZYMES
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Quarles & Brady
;     STREET: 411 East Wisconsin Avenue
;     CITY: Milwaukee
;     STATE: Wisconsin
;     COUNTRY: U.S.A.
;     ZIP: 53202-4497
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US96/00728
;     FILING DATE:
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Baker, Jean C.
```

```
     REGISTRATION NUMBER: 35,433
;    REFERENCE/DOCKET NUMBER: 660336.90489
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (414) 277-5709
;      TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 249 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;    MOLECULE TYPE: DNA (genomic)
PCT-US96-00728-2

Query Match          9.4%;  Score 17;  DB 5;  Length 249;
Best Local Similarity 100.0%;  Pred. No. 5.6;
Matches  17;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       61 ATGGACAATTGCGCGCCA 77
            |||||||||||||||||
Db       63 ATGGACAATTGCGCGCCA 47

RESULT 6
US-08-446-794A-3
; Sequence 3, Application US/08446794A
; Patent No. 5747327
; GENERAL INFORMATION:
;   APPLICANT: UEKI, JUN
;   APPLICANT: MORIOKA, SHINJI
;   TITLE OF INVENTION: PHOSPHOLIPASE D GENE ORIGINATED FROM
;   TITLE OF INVENTION: PLANT
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
;     STREET: P.O. BOX 747
;     CITY: FALLS CHURCH
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22040-0747
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/446,794A
;     FILING DATE:
;     CLASSIFICATION: 536
;   ATTORNEY/AGENT INFORMATION:
;     NAME: MURPHY JR, GERALD M
;     REGISTRATION NUMBER: 28,977
;     REFERENCE/DOCKET NUMBER: 0760-0203P
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-205-8000
;     TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2804 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 107..2542
US-08-446-794A-3

Query Match          9.4%;  Score 17;  DB 1;  Length 2804;
Best Local Similarity 100.0%;  Pred. No. 5.8;
Matches  17;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       24 GCTCGAAGACGTCTTCC 40
            |||||||||||||||||
Db     2290 GCTCGAAGACGTCTTCC 2306

RESULT 7
US-09-328-111-381/c
; Sequence 381, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
;   APPLICANT: Endege, Wilson O.
;   APPLICANT: Steinmann, Kathleen E.
;   APPLICANT: Astle, Jon H.
;   APPLICANT: Burgess, Christopher C.
;   APPLICANT: Bushnell, Steven E.
;   APPLICANT: Carroll III, Eddie
;   APPLICANT: Catino, Theodore J.
;   APPLICANT: Derti, Adnan
;   APPLICANT: Ford, Donna M.
;   APPLICANT: Lewis, Marcia E.
;   APPLICANT: Monahan, John E.
;   APPLICANT: Schlegel, Robert
;   TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;   TITLE OF INVENTION: PRODUCTS
;   FILE REFERENCE: CCD-257 (US)
;   CURRENT APPLICATION NUMBER: US/09/328,111
;   CURRENT FILING DATE: 1999-06-08
;   EARLIER APPLICATION NUMBER: US 60/088,801
;   EARLIER FILING DATE: 1998-06-10
;   NUMBER OF SEQ ID NOS: 850
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 381
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)...(622)
;   OTHER INFORMATION: n = A,T,C or G
US-09-328-111-381

Query Match          8.8%;  Score 16;  DB 4;  Length 622;
Best Local Similarity 100.0%;  Pred. No. 19;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       57 TTATATGGACAATTGG 72
            ||||||||||||||||
Db       22 TTATATGGACAATTGG 7

RESULT 8
US-09-342-648-7/c
; Sequence 7, Application US/09342648
; Patent No. 6248584
; GENERAL INFORMATION:
;   APPLICANT: Cahoon, Rebecca E.
;   APPLICANT: Odell, Joan
;   APPLICANT: Rafalski, Antoni
;   TITLE OF INVENTION: Transcription Coactivators
;   FILE REFERENCE: BB-1169-B
;   CURRENT APPLICATION NUMBER: US/09/342,648
;   CURRENT FILING DATE: 1999-06-29
;   EARLIER APPLICATION NUMBER: 60/092,659
;   EARLIER FILING DATE: July 13, 1998
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 1626
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;     NAME/KEY: unsure
;     LOCATION: (1621)
```

US-09-342-648-7

Query Match   8.8%; Score 16; DB 4; Length 1626;
Best Local Similarity   100.0%; Pred. No. 20;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   85 ACAGCTGAGCAAGAGG 100
          ||||||||||||||||
Db   1166 ACAGCTGAGCAAGAGG 1151

RESULT 9
US-08-975-080-35
; Sequence 35, Application US/08975080
; Patent No. 6245523
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,080
; FILING DATE: 20-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,435
; FILING DATE: 20-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match   8.8%; Score 16; DB 4; Length 14796;
Best Local Similarity   100.0%; Pred. No. 20;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTATTACTTATA 10824

RESULT 10
US-09-630-706-10
; Sequence 10, Application US/09630706
; Patent No. 6277640
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
; FILE REFERENCE: RTS-0053
; CURRENT APPLICATION NUMBER: US/09/630,706

; CURRENT FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 94
; SEQ ID NO 10
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-09-630-706-10

Query Match   8.8%; Score 16; DB 4; Length 14796;
Best Local Similarity   100.0%; Pred. No. 20;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTATTACTTATA 10824

RESULT 11
US-09-496-694B-3
; Sequence 3, Application US/09496694B
; Patent No. 6335194
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0439
; CURRENT APPLICATION NUMBER: US/09/496,694B
; CURRENT FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 3
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match   8.8%; Score 16; DB 4; Length 14796;
Best Local Similarity   100.0%; Pred. No. 20;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTATTACTTATA 10824

RESULT 12
US-09-221-017B-994
; Sequence 994, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:

```
   APPLICANT: Ross, Bruce C.
   TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
   NUMBER OF SEQUENCES: 1120
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: MORRISON & FOERSTER
      STREET: 755 PAGE MILL ROAD
      CITY: Palo Alto
      STATE: CA
      COUNTRY: USA
      ZIP: 94304-1018
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette
      COMPUTER: IBM Compatible
      OPERATING SYSTEM: Windows
      SOFTWARE: FastSEQ for Windows Version 2.0b
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/221,017B
      FILING DATE: 23-DEC-1998
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: PP1182
      FILING DATE: 31-DEC-1997
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: PP1546
      FILING DATE: 30-JAN-1998
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: PP2911
      FILING DATE: 09-APR-1998
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: PCT/AU98/01023
      FILING DATE: 10-DEC-1998
   ATTORNEY/AGENT INFORMATION:
      NAME: Monroy, Gladys H
      REGISTRATION NUMBER: 32,430
      REFERENCE/DOCKET NUMBER: 27340-20021.00
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 650-813-5600
      TELEFAX: 650-494-0792
      TELEX: 706141
   INFORMATION FOR SEQ ID NO: 994:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 797 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: circular
   MOLECULE TYPE: DNA (genomic)
   HYPOTHETICAL: NO
   ANTI-SENSE: UNKNOWN
   ORIGINAL SOURCE:
      ORGANISM: PORYPHYROMONAS GINGIVALIS
   FEATURE:
      NAME/KEY: misc_feature
      LOCATION: 1..797
US-09-221-017B-994

Query Match    8.3%;  Score 15;  DB 4;  Length 797;
Best Local Similarity  100.0%;  Pred. No. 66;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    101 CCCTCCAAGCCAAAG 115
          |||||||||||||||
Db    422 CCCTCCAAGCCAAAG 436


RESULT 13
US-09-134-001C-1175
; Sequence 1175, Application US/09134001C
; Patent No. 6380370
GENERAL INFORMATION:
;   APPLICANT: Lynn Doucette-Stamm et al
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;   TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: GTC-007
;   CURRENT APPLICATION NUMBER: US/09/134,001C
;   CURRENT FILING DATE: 1998-08-13
;   PRIOR APPLICATION NUMBER: US 60/064,964
;   PRIOR FILING DATE: 1997-11-08
;   PRIOR APPLICATION NUMBER: US 60/055,779
;   PRIOR FILING DATE: 1997-08-14
;   NUMBER OF SEQ ID NOS: 5674
;   SEQ ID NO 1175
;   LENGTH: 1854
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1175

Query Match    8.3%;  Score 15;  DB 4;  Length 1854;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    147 GTACCTCATGGTGAT 161
          |||||||||||||||
Db   1750 GTACCTCATGGTGAT 1764


RESULT 14
US-08-961-527-174/C
; Sequence 174, Application US/08961527
; Patent No. 6420135
GENERAL INFORMATION:
;   APPLICANT: Charles Kunsch
;   TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;   NUMBER OF SEQUENCES: 391
;   CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Human Genome Sciences, Inc.
;      STREET: 9410 Key West Avenue
;      CITY: Rockville
;      STATE: Maryland
;      COUNTRY: USA
;      ZIP: 20850
;   COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER: HP Vectra 486/33
;      OPERATING SYSTEM: MSDOS version 6.2
;      SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/961,527
;      FILING DATE:
;      CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;      NAME: Brookes, A. Anders
;      REGISTRATION NUMBER: 36,373
;      REFERENCE/DOCKET NUMBER: PB340P1
;   TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (301) 309-8504
;      TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 174:
;   SEQUENCE CHARACTERISTICS:
;      LENGTH: 2069 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
US-08-961-527-174

Query Match    8.3%;  Score 15;  DB 4;  Length 2069;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    161 TGATTGGAATGTTCT 175
          |||||||||||||||
Db    272 TGATTGGAATGTTCT 258
```

RESULT 15
US-08-804-439A-12/c
; Sequence 12, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
;   APPLICANT: Rose, Timothy M.
;   APPLICANT: Bosch, Marnix L.
;   APPLICANT: Strand, Kurt
;   TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
;   TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
;   NUMBER OF SEQUENCES: 113
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson P.C.
;   STREET: 4225 Executive Square, Ste 1400
;   CITY: La Jolla
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92037
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/804,439A
;   FILING DATE: February 21, 1997
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Haile, Lisa A.
;   REGISTRATION NUMBER: 38,347
;   REFERENCE/DOCKET NUMBER: 09176/004001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 678-5070
;   TELEFAX: (619) 678-5099
;   TELEX:
US-08-804-439A-12

Query Match   8.3%; Score 15; DB 3; Length 2608;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   163 ATTGGAATGTTCTCT 177
         |||||||||||||||
Db   768 ATTGGAATGTTCTCT 754

RESULT 16
US-08-720-229-12/c
; Sequence 12, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
;   APPLICANT: Rose, Timothy M.
;   APPLICANT: Bosch, Marnix L.
;   APPLICANT: Strand, Kurt
;   TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
;   TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
;   NUMBER OF SEQUENCES: 100
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Morrison & Foerster
;   STREET: 755 Page Mill Road
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304-1018
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/720,229
;   FILING DATE: 26-SEP-1996
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Schiff, J. Michael
;   REGISTRATION NUMBER: 40,253
;   REFERENCE/DOCKET NUMBER: 29938-20002.00
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 813-5600
;   TELEFAX: (415) 494-0792
;   TELEX: 706141
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
    LENGTH: 2608 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
US-08-720-229-12

Query Match   8.3%; Score 15; DB 3; Length 2608;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   163 ATTGGAATGTTCTCT 177
         |||||||||||||||
Db   768 ATTGGAATGTTCTCT 754

RESULT 17
US-08-729-955A-2
; Sequence 2, Application US/08729955A
; Patent No. 5932417
; GENERAL INFORMATION:
;   APPLICANT: Birnbaumer, Lutz
;   APPLICANT: Zhu, Xi
;   TITLE OF INVENTION: Method And Compounds For Controlling
;   TITLE OF INVENTION: Capacitative Calcium Ion Entry Into Mammalian Cells
;   TITLE OF INVENTION: Essential for Agonist-Activated Capacitative Ca2+
;   TITLE OF INVENTION: Entry
;   NUMBER OF SEQUENCES: 32
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Poms, Smith, Lande & Rose
;   STREET: 2029 Century Park East, Suite 3800
;   CITY: Los Angeles
;   STATE: California
;   COUNTRY: USA
;   ZIP: 90067
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/729,955A
;   FILING DATE: October 15, 1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/025,111
;   FILING DATE: August 29, 1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Oldenkamp, David J.
;   REGISTRATION NUMBER: 29,421
;   REFERENCE/DOCKET NUMBER: 120186
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (310) 788-5000
;   TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
    LENGTH: 3258 base pairs
    TYPE: nucleic acid

```
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: DNA
;  HYPOTHETICAL: NO
;  ANTI-SENSE: NO
;  ORIGINAL SOURCE:
;    ORGANISM:
;    INDIVIDUAL ISOLATE: Mtrp4
US-08-729-955A-2

Query Match    8.3%;   Score 15;   DB 2;   Length 3258;
Best Local Similarity  100.0%;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    160 ATGATTGGAATGTTC 174
          ||||||||||||||
Db   2232 ATGATTGGAATGTTC 2246

RESULT 18
US-08-726-214-7
;  Sequence 7, Application US/08726214
;  Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT: Tang, Wei-Jen
;    APPLICANT: Gilman, Alfred G.
;    TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION: AND USES THEREFOR
;    NUMBER OF SEQUENCES: 31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Arnold, White & Durkee
;      STREET: P.O. Box 4433
;      CITY: Houston
;      STATE: Texas
;      COUNTRY: United States of America
;      ZIP: 77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/726,214
;      FILING DATE: Concurrently Herewith
;      CLASSIFICATION: 435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 60/005,498
;      FILING DATE: 04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Highlander, Steven L.
;      REGISTRATION NUMBER: 37,642
;      REFERENCE/DOCKET NUMBER: UTSD:450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (512) 418-3000
;      TELEFAX: (512) 474-7577
;  INFORMATION FOR SEQ ID NO: 7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 3357 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
US-08-726-214-7

Query Match    8.3%;   Score 15;   DB 3;   Length 3357;
Best Local Similarity  100.0%;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    103 CTCCAAGCCAAAGTT 117
          |||||||||||||||
Db   2815 CTCCAAGCCAAAGTT 2829

RESULT 19
```

```
US-08-961-527-143/c
;  Sequence 143, Application US/08961527
;  Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT: Charles Kunsch
;    TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES: 391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Human Genome Sciences, Inc.
;      STREET: 9410 Key West Avenue
;      CITY: Rockville
;      STATE: Maryland
;      COUNTRY: USA
;      ZIP: 20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER: HP Vectra 486/33
;      OPERATING SYSTEM: MSDOS version 6.2
;      SOFTWARE: ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/961,527
;      FILING DATE:
;      CLASSIFICATION: 424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Brookes, A. Anders
;      REGISTRATION NUMBER: 36,373
;      REFERENCE/DOCKET NUMBER: PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (301) 309-8504
;      TELEFAX: (301) 309-8512
;  INFORMATION FOR SEQ ID NO: 143:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 4965 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
US-08-961-527-143

Query Match    8.3%;   Score 15;   DB 4;   Length 4965;
Best Local Similarity  100.0%;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    161 TGATTGGAATGTTCT 175
          |||||||||||||||
Db   4533 TGATTGGAATGTTCT 4519

RESULT 20
US-08-961-527-137
;  Sequence 137, Application US/08961527
;  Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT: Charles Kunsch
;    TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES: 391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Human Genome Sciences, Inc.
;      STREET: 9410 Key West Avenue
;      CITY: Rockville
;      STATE: Maryland
;      COUNTRY: USA
;      ZIP: 20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER: HP Vectra 486/33
;      OPERATING SYSTEM: MSDOS version 6.2
;      SOFTWARE: ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/961,527
;      FILING DATE:
```

```
                                              Qy        1 ACTTTATCCAATTC 15
                                                          ||||||||||||||
                                              Db     8919 ACTTTATCCAATTC 8905

                                              RESULT 22
                                              US-08-458-434A-6/c
                                              ; Sequence 6, Application US/08458434A
                                              ; Patent No. 6083690
                                              ; GENERAL INFORMATION:
                                              ;   APPLICANT: Harris Ph.D., Stephen E.
                                              ;   APPLICANT: Mundy M.D., Gregory R.
                                              ;   APPLICANT: Gosh-Choudhury Ph.D., Nandini
                                              ;   APPLICANT: Feng Ph.D., Jian Q.
                                              ;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING
                                              ;   TITLE OF INVENTION: OSTEOGENIC AGENTS
                                              ;   NUMBER OF SEQUENCES: 13
                                              ;   CORRESPONDENCE ADDRESS:
                                              ;     ADDRESSEE: James C. Weseman, Esq.
                                              ;     STREET: 401 B. Street, Suite 1700
                                              ;     CITY: San Diego
                                              ;     STATE: CA
                                              ;     COUNTRY: USA
                                              ;     ZIP: 92101
                                              ;   COMPUTER READABLE FORM:
                                              ;     MEDIUM TYPE: Floppy disk
                                              ;     COMPUTER: IBM PC compatible
                                              ;     OPERATING SYSTEM: PC-DOS/MS-DOS
                                              ;     SOFTWARE: PatentIn Release #1.0, Version #1.25
                                              ;   CURRENT APPLICATION DATA:
                                              ;     APPLICATION NUMBER: US/08/458,434A
                                              ;     FILING DATE:
                                              ;   CLASSIFICATION: 435
                                              ;   ATTORNEY/AGENT INFORMATION:
                                              ;     NAME: Weseman, James C.
                                              ;     REGISTRATION NUMBER: 30,507
                                              ;     REFERENCE/DOCKET NUMBER: P00060US0
                                              ;   TELECOMMUNICATION INFORMATION:
                                              ;     TELEPHONE: (619) 699-3604
                                              ;     TELEFAX: 619-236-1048
                                              ; INFORMATION FOR SEQ ID NO:   6:
                                              ;   SEQUENCE CHARACTERISTICS:
                                              ;     LENGTH: 15144 base pairs
                                              ;     TYPE: nucleic acid
                                              ;     STRANDEDNESS: single
                                              ;     TOPOLOGY: linear
                                              ;     MOLECULE TYPE: DNA (genomic)
                                              US-08-458-434A-6

                                              Query Match             8.3%;  Score 15;  DB 3;  Length 15144;
                                              Best Local Similarity  100.0%;  Pred. No. 69;
                                              Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

                                              Qy       31 GACGTCTTCCGAAGG 45
                                                          ||||||||||||||
                                              Db     4226 GACGTCTTCCGAAGG 4212

                                              RESULT 23
                                              US-08-735-609-4
                                              ; Sequence 4, Application US/08735609
                                              ; Patent No. 5955360
                                              ; GENERAL INFORMATION:
                                              ;   APPLICANT: Chamberlain, Jeffrey S.
                                              ;   APPLICANT: Amalfitano, Andrea
                                              ;   APPLICANT: Hauser, Michael A.
                                              ;   APPLICANT: Kumar-Singh, Rajendra
                                              ;   APPLICANT: Hartigan-O'Connor, Dennis J.
                                              ;   TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
                                              ;   NUMBER OF SEQUENCES: 15
                                              ;   CORRESPONDENCE ADDRESS:
                                              ;     ADDRESSEE: Medlen & Carroll, LLP
```

```
;   CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 137:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 12666 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-961-527-137

Query Match             8.3%;  Score 15;  DB 4;  Length 12666;
Best Local Similarity  100.0%;  Pred. No. 69;
Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      156 GGTGATGATTGGAAT 170
             ||||||||||||||
Db    10747 GGTGATGATTGGAAT 10761

RESULT 21
US-08-961-527-128/c
; Sequence 128, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
;   APPLICANT: Charles Kunsch
;   TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;   NUMBER OF SEQUENCES: 391
;   CORRESPONDENCE ADDRESS:
;     ADDRESSE: Human Genome Sciences, Inc.
;     STREET: 9410 Key West Avenue
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER: HP Vectra 486/33
;     OPERATING SYSTEM: MSDOS version 6.2
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/961,527
;     FILING DATE:
;   CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 128:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 13440 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-961-527-128

Query Match             8.3%;  Score 15;  DB 4;  Length 13440;
Best Local Similarity  100.0%;  Pred. No. 69;
Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
            STREET: 220 Montgomery Street, Suite 2200
            CITY: San Francisco
            STATE: California
            COUNTRY: United States Of America
            ZIP: 94104
       COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.30
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/08/735,609
            FILING DATE:
            CLASSIFICATION: 435
       ATTORNEY/AGENT INFORMATION:
            NAME: Ingolia, Diane E.
            REGISTRATION NUMBER: 40,027
            REFERENCE/DOCKET NUMBER: UM-02484
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (415) 705-8410
            TELEFAX: (415) 397-8338
       INFORMATION FOR SEQ ID NO: 4:
            SEQUENCE CHARACTERISTICS:
                 LENGTH: 34303 base pairs
                 TYPE: nucleic acid
                 STRANDEDNESS: double
                 TOPOLOGY: linear
            MOLECULE TYPE: other nucleic acid
                 DESCRIPTION: /desc = "DNA"
US-08-735-609-4

Query Match             8.3%;   Score 15;  DB 2;  Length 34303;
Best Local Similarity  100.0%;  Pred. No. 70;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     21 GACGCTGGAAGACGT 35
          ||||||||||||||
Db   8439 GACGCTGGAAGACGT 8453

RESULT 24
US-08-735-609-4
; Sequence 4, Application US/08735609
; Patent No. 5594132
; GENERAL INFORMATION:
       APPLICANT: Chamberlain, Jeffrey S.
                  Amalfitano, Andrea
                  Hauser, Michael A.
                  Kumar-Singh, Rajendra
                  Hartigan-O'Connor, Dennis J.
       TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
       NUMBER OF SEQUENCES: 15
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Medlen & Carroll, LLP
            STREET: 220 Montgomery Street, Suite 2200
            CITY: San Francisco
            STATE: California
            COUNTRY: United States Of America
            ZIP: 94104
       COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.30
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/08/735,609
            FILING DATE: 23-Oct-1996
            CLASSIFICATION: <Unknown>
       ATTORNEY/AGENT INFORMATION:
            NAME: Ingolia, Diane E.
            REGISTRATION NUMBER: 40,027
            REFERENCE/DOCKET NUMBER: UM-02484
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (415) 705-8410
            TELEFAX: (415) 397-8338
       INFORMATION FOR SEQ ID NO: 4:
            SEQUENCE CHARACTERISTICS:
                 LENGTH: 34303 base pairs
                 TYPE: nucleic acid
                 STRANDEDNESS: double
                 TOPOLOGY: linear
            MOLECULE TYPE: other nucleic acid
                 DESCRIPTION: /desc = "DNA"
                 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-735-609-4

Query Match             8.3%;   Score 15;  DB 2;  Length 34303;
Best Local Similarity  100.0%;  Pred. No. 70;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     21 GACGCTGGAAGACGT 35
          ||||||||||||||
Db   8439 GACGCTGGAAGACGT 8453

RESULT 25
US-09-315-372-4
; Sequence 4, Application US/09315372
; Patent No. 6057158
; GENERAL INFORMATION:
       APPLICANT: Chamberlain, Jeffrey S.
       APPLICANT: Amalfitano, Andrea
       APPLICANT: Hauser, Michael A.
       APPLICANT: Kumar-Singh, Rajendra
       APPLICANT: Hartigan-O'Connor, Dennis J.
       TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
       NUMBER OF SEQUENCES: 15
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Medlen & Carroll, LLP
            STREET: 220 Montgomery Street, Suite 2200
            CITY: San Francisco
            STATE: California
            COUNTRY: United States Of America
            ZIP: 94104
       COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.30
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/09/315,372
            FILING DATE:
            CLASSIFICATION:
       PRIOR APPLICATION DATA:
            APPLICATION NUMBER: US 08/735,609
            FILING DATE:
       ATTORNEY/AGENT INFORMATION:
            NAME: Ingolia, Diane E.
            REGISTRATION NUMBER: 40,027
            REFERENCE/DOCKET NUMBER: UM-02484
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (415) 705-8410
            TELEFAX: (415) 397-8338
       INFORMATION FOR SEQ ID NO: 4:
            SEQUENCE CHARACTERISTICS:
                 LENGTH: 34303 base pairs
                 TYPE: nucleic acid
                 STRANDEDNESS: double
                 TOPOLOGY: linear
            MOLECULE TYPE: other nucleic acid
                 DESCRIPTION: /desc = "DNA"
US-09-315-372-4

Query Match             8.3%;   Score 15;  DB 3;  Length 34303;
Best Local Similarity  100.0%;  Pred. No. 70;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy     21 GACGCTGGAAGACGT 35
          |||||||||||||||
Db   8439 GACGCTGGAAGACGT 8453

RESULT 26
US-09-244-752-4
; Sequence 4, Application US/09244752
; Patent No. 6063622
; GENERAL INFORMATION:
;    APPLICANT: Chamberlain, Jeffrey S.
;    APPLICANT: Amalfitano, Andrea
;    APPLICANT: Hauser, Michael A.
;    APPLICANT: Kumar-Singh, Rajendra
;    APPLICANT: Hartigan-O'Connor, Dennis J.
;    TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;    NUMBER OF SEQUENCES: 15
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Medlen & Carroll, LLP
;       STREET: 220 Montgomery Street, Suite 2200
;       CITY: San Francisco
;       STATE: California
;       COUNTRY: United States Of America
;       ZIP: 94104
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/244,752
;       FILING DATE:
;       CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/735,609
;       FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Ingolia, Diane E.
;       REGISTRATION NUMBER: 40,027
;       REFERENCE/DOCKET NUMBER: UM-02484
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (415) 705-8410
;       TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 34303 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;    MOLECULE TYPE: other nucleic acid
;       DESCRIPTION: /desc = "DNA"
US-09-244-752-4

Query Match         8.3%;  Score 15;  DB 3;  Length 34303;
Best Local Similarity  100.0%;  Pred. No. 70;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     21 GACGCTGGAAGACGT 35
          |||||||||||||||
Db   8439 GACGCTGGAAGACGT 8453

RESULT 27
US-09-245-497-4
; Sequence 4, Application US/09245497
; Patent No. 6083750
; GENERAL INFORMATION:
;    APPLICANT: Chamberlain, Jeffrey S.
;    APPLICANT: Amalfitano, Andrea
;    APPLICANT: Hauser, Michael A.
;    APPLICANT: Kumar-Singh, Rajendra
;    APPLICANT: Hartigan-O'Connor, Dennis J.
;    TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;    NUMBER OF SEQUENCES: 15
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Medlen & Carroll, LLP
;       STREET: 220 Montgomery Street, Suite 2200
;       CITY: San Francisco
;       STATE: California
;       COUNTRY: United States Of America
;       ZIP: 94104
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/245,497
;       FILING DATE:
;       CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/735,609
;       APPLICATION NUMBER: <B> FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Ingolia, Diane E.
;       REGISTRATION NUMBER: 40,027
;       REFERENCE/DOCKET NUMBER: UM-02484
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (415) 705-8410
;       TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 34303 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;    MOLECULE TYPE: other nucleic acid
;       DESCRIPTION: /desc = "DNA"
US-09-245-497-4

Query Match         8.3%;  Score 15;  DB 3;  Length 34303;
Best Local Similarity  100.0%;  Pred. No. 70;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     21 GACGCTGGAAGACGT 35
          |||||||||||||||
Db   8439 GACGCTGGAAGACGT 8453

RESULT 28
US-09-562-919-4
; Sequence 4, Application US/09562919
; Patent No. 6451596
; GENERAL INFORMATION:
;    APPLICANT: Chamberlain, Jeffrey S.
;    APPLICANT: Amalfitano, Andrea
;    APPLICANT: Hauser, Michael A.
;    APPLICANT: Kumar-Singh, Rajendra
;    APPLICANT: Hartigan-O'Connor, Dennis J.
;    TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;    NUMBER OF SEQUENCES: 15
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Medlen & Carroll, LLP
;       STREET: 220 Montgomery Street, Suite 2200
;       CITY: San Francisco
;       STATE: California
;       COUNTRY: United States Of America
;       ZIP: 94104
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/562,919
```

```
                FILING DATE: 02-May-2000
                CLASSIFICATION: <Unknown>
            PRIOR APPLICATION DATA:
                APPLICATION NUMBER: US/08/735,609
                FILING DATE: 23-Oct-1996
            ATTORNEY/AGENT INFORMATION:
                NAME: Ingolia, Diane E.
                REGISTRATION NUMBER: 40,027
                REFERENCE/DOCKET NUMBER: UM-02484
            TELECOMMUNICATION INFORMATION:
                TELEPHONE: (415) 705-8410
                TELEFAX: (415) 397-8338
        INFORMATION FOR SEQ ID NO: 4:
            SEQUENCE CHARACTERISTICS:
                LENGTH: 34303 base pairs
                TYPE: nucleic acid
                STRANDEDNESS: double
                TOPOLOGY: linear
                MOLECULE TYPE: other nucleic acid
                DESCRIPTION:/desc = "DNA"
            SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-562-919-4

Query Match            8.3%;  Score 15;  DB 4;  Length 34303;
Best Local Similarity 100.0%;  Pred. No. 70;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        21 GACGCTGGAAGACGT 35
             |||||||||||||||
Db      8439 GACGCTGGAAGACGT 8453

RESULT 29
US-08-374-483-6
; Sequence 6, Application US/08374483
; Patent No. 5880102
    GENERAL INFORMATION:
        APPLICANT: GEORGE, SAMUEL E.
        APPLICANT: BLAZING, MICHAEL A.
        TITLE OF INVENTION: ADENOVIRAL VECTOR SYSTEM
        NUMBER OF SEQUENCES: 9
        CORRESPONDENCE ADDRESS:
            ADDRESSEE: NIXON & VANDERHYE P.C.
            STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
            CITY: ARLINGTON
            STATE: VIRGINIA
            COUNTRY: U.S.A.
            ZIP: 22201-4714
        COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.25
        CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/08/374,483
            FILING DATE: 17-JAN-1995
            CLASSIFICATION:
        ATTORNEY/AGENT INFORMATION:
            NAME: WILSON, MARY J.
            REGISTRATION NUMBER: 32,955
            REFERENCE/DOCKET NUMBER: 1579-83
        TELECOMMUNICATION INFORMATION:
            TELEPHONE: (703) 816-4000
            TELEFAX: (703) 816-4100
            TELEX: 200797 NIXN UR
    INFORMATION FOR SEQ ID NO: 6:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 34382 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
            MOLECULE TYPE: DNA (genomic)
US-08-374-483-6


Query Match            8.3%;  Score 15;  DB 2;  Length 34382;
Best Local Similarity 100.0%;  Pred. No. 70;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        21 GACGCTGGAAGACGT 35
             |||||||||||||||
Db      7859 GACGCTGGAAGACGT 7873

RESULT 30
US-08-973-334-3
; Sequence 3, Application US/08973334
; Patent No. 6261551
    GENERAL INFORMATION:
        APPLICANT: Wilson, James M.
        APPLICANT: Fisher, Krishna J.
        APPLICANT: Gao, Guang-Ping
        TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
        TITLE OF INVENTION: Associated virus, Cell Lines, and
        TITLE OF INVENTION: Methods of Production and Use
        TITLE OF INVENTION: Thereof
        NUMBER OF SEQUENCES: 5
        CORRESPONDENCE ADDRESS:
            ADDRESSEE: Howson and Howson
            STREET: Box 457, 321 No. 6261551ristown Road
            CITY: Spring House
            STATE: PA
            COUNTRY: USA
            ZIP: 19477
        COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release 1.0 Version 1.30
        CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/08/973,334
            FILING DATE:
            CLASSIFICATION: 514
        PRIOR APPLICATION DATA:
            APPLICATION NUMBER: US 08/462,014
            FILING DATE: 05-JUN-1995
        PRIOR APPLICATION DATA:
            APPLICATION NUMBER: US 08/549,489
            FILING DATE: 27-OCT-1995
        ATTORNEY/AGENT INFORMATION:
            NAME: Bak, Mary E.
            REGISTRATION NUMBER: 31,215
            REFERENCE/DOCKET NUMBER: GNVPN012CIPUSA
        TELECOMMUNICATION INFORMATION:
            TELEPHONE: (215) 540-9206
            TELEFAX: (215) 540-5818
    INFORMATION FOR SEQ ID NO: 3:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 35408 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: not relevant
            TOPOLOGY: not relevant
            MOLECULE TYPE: other nucleic acid
US-08-973-334-3

Query Match            8.3%;  Score 15;  DB 4;  Length 35408;
Best Local Similarity 100.0%;  Pred. No. 70;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        21 GACGCTGGAAGACGT 35
             |||||||||||||||
Db      8264 GACGCTGGAAGACGT 8278

RESULT 31
US-09-563-869A-3
; Sequence 3, Application US/09563869A
```

Patent No. 6270996
GENERAL INFORMATION:
  APPLICANT: Wilson, James M.
             Fisher, Krishna J.
             Gao, Guang-Ping
  TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
                      Associated virus, Cell Lines, and
                      Methods of Production and Use
                      Thereof
  NUMBER OF SEQUENCES: 5
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Howson and Howson
    STREET: Box 457, 321 No. 6270996ristown Road
    CITY: Spring House
    STATE: PA
    COUNTRY: USA
    ZIP: 19477
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release 1.0 Version 1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/563,869A
    FILING DATE: 03-May-2000
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/973,334
    FILING DATE: <Unknown>
    APPLICATION NUMBER: US 08/549,489
    FILING DATE: 27-OCT-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: Bak, Mary E.
    REGISTRATION NUMBER: 31,215
    REFERENCE/DOCKET NUMBER: GNVPN012CIPUSA
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (215) 540-9206
    TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 35408 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: not relevant
    TOPOLOGY: not relevant
    MOLECULE TYPE: other nucleic acid
  SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-563-869A-3

Query Match   8.3%; Score 15; DB 4; Length 35408;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   21 GACGCTGGAAGACGT 35
         ||||||||||||||
Db 8264 GACGCTGGAAGACGT 8278

RESULT 32
US-08-549-489-3
; Sequence 3, Application US/08549489
; Patent No. 6281010
GENERAL INFORMATION:
  APPLICANT: Wilson, James M.
  APPLICANT: Gao, Guang-Ping
  TITLE OF INVENTION: No. 62810101el Adenovirus Gene Therapy Vehicle
  TITLE OF INVENTION: and Cell Line
  NUMBER OF SEQUENCES: 3
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Howson and Howson
    STREET: Box 457, 321 No. 6281010ristown Road
    CITY: Spring House
    STATE: PA
    COUNTRY: USA
    ZIP: 19477
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/549,489
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/462,014
    FILING DATE: 08-JUN-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: Bak, Mary E.
    REGISTRATION NUMBER: 31,215
    REFERENCE/DOCKET NUMBER: GNVPN013
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (215) 540-9206
    TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 35408 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: not relevant
    TOPOLOGY: not relevant
    MOLECULE TYPE: other nucleic acid
US-08-549-489-3

Query Match   8.3%; Score 15; DB 4; Length 35408;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   21 GACGCTGGAAGACGT 35
         ||||||||||||||
Db 8264 GACGCTGGAAGACGT 8278

RESULT 33
US-08-735-609-1
; Sequence 1, Application US/08735609
; Patent No. 5955360
GENERAL INFORMATION:
  APPLICANT: Chamberlain, Jeffrey S.
  APPLICANT: Amalfitano, Andrea
  APPLICANT: Hauser, Michael A.
  APPLICANT: Kumar-Singh, Rajendra
  APPLICANT: Hartigan-O'Connor, Dennis J.
  TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
  NUMBER OF SEQUENCES: 15
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Medlen & Carroll, LLP
    STREET: 220 Montgomery Street, Suite 2200
    CITY: San Francisco
    STATE: California
    COUNTRY: United States Of America
    ZIP: 94104
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/735,609
    FILING DATE:
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Ingolia, Diane E.
    REGISTRATION NUMBER: 40,027
    REFERENCE/DOCKET NUMBER: UM-02484
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415) 705-8410
    TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-735-609-1

Query Match   8.3%;   Score 15;   DB 2;   Length 35935;
Best Local Similarity 100.0%;   Pred. No. 70;
Matches   15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   21 GACGCTGGAAGACGT 35
            ||||||||||||||
Db   6936 GACGCTGGAAGACGT 6950

RESULT 34
US-08-735-609-1
; Sequence 1, Application US/08735609
; Patent No. 5994132
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
;   Amalfitano, Andrea
;   Hauser, Michael A.
;   Kumar-Singh, Rajendra
;   Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Medlen & Carroll, LLP
;   STREET: 220 Montgomery Street, Suite 2200
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: United States Of America
;   ZIP: 94104
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/735,609
;   FILING DATE: 23-Oct-1996
;   CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Ingolia, Diane E.
;   REGISTRATION NUMBER: 40,027
;   REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 705-8410
;   TELEFAX: (415) 397-8338
;   DESCRIPTION: /desc = "DNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-735-609-1

Query Match   8.3%;   Score 15;   DB 2;   Length 35935;
Best Local Similarity 100.0%;   Pred. No. 70;
Matches   15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   21 GACGCTGGAAGACGT 35
            ||||||||||||||
Db   6936 GACGCTGGAAGACGT 6950

RESULT 35
US-08-379-452-43
; Sequence 43, Application US/08379452
; Patent No. 6040174
; GENERAL INFORMATION:
; APPLICANT: IMLER, Jean-Luc
; APPLICANT: MEHTALI, Majid
; APPLICANT: PAVIRANI, Andrea
; TITLE OF INVENTION: DEFECTIVE ADENOVIRUSES AND CORRESPONDING
; TITLE OF INVENTION: COMPLEMENTATION LINES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;   STREET: 1737 King Street, Suite 500
;   CITY: Alexandria
;   STATE: Virginia
;   COUNTRY: United States
;   ZIP: 22314-2756
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/379,452
;   FILING DATE: 26-JAN-1995
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: WO PCT/FR94/00624
;   FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 93 06482
;   FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
;   NAME: Dadio, Susan M.
;   REGISTRATION NUMBER: 40,373
;   REFERENCE/DOCKET NUMBER: 029395-002
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 35935 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-379-452-43

Query Match   8.3%;   Score 15;   DB 3;   Length 35935;
Best Local Similarity 100.0%;   Pred. No. 70;
Matches   15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   21 GACGCTGGAAGACGT 35
            ||||||||||||||
Db   6936 GACGCTGGAAGACGT 6950

RESULT 36
US-09-315-372-1
; Sequence 1, Application US/09315372
; Patent No. 6057158
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Amalfitano, Andrea
; APPLICANT: Hauser, Michael A.
; APPLICANT: Kumar-Singh, Rajendra
; APPLICANT: Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
;   ADDRESSE: Medlen & Carroll, LLP
;   STREET: 220 Montgomery Street, Suite 2200
;   CITY: San Francisco
;   STATE: California

```
          COUNTRY: United States of America
          ZIP: 94104
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/09/315,372
          FILING DATE:
          CLASSIFICATION:
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/735,609
          FILING DATE:
     ATTORNEY/AGENT INFORMATION:
          NAME: Ingolia, Diane E.
          REGISTRATION NUMBER: 40,027
          REFERENCE/DOCKET NUMBER: UM-02484
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: (415) 705-8410
          TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 35935 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
     MOLECULE TYPE: other nucleic acid
          DESCRIPTION: /desc = "DNA"
US-09-315-372-1

Query Match      8.3%; Score 15; DB 3; Length 35935;
Best Local Similarity 100.0%; Pred. No. 70;
Matches  15; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy     21 GACGCTGGAAGACGT 35
          ||||||||||||||
Db   6936 GACGCTGGAAGACGT 6950

RESULT 37
US-09-244-752-1
; Sequence 1, Application US/09244752
; Patent No. 6063622
; GENERAL INFORMATION:
;     APPLICANT: Chamberlain, Jeffrey S.
;     APPLICANT: Amalfitano, Andrea
;     APPLICANT: Hauser, Michael A.
;     APPLICANT: Kumar-Singh, Rajendra
;     APPLICANT: Hartigan-O'Connor, Dennis J.
;     TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;     NUMBER OF SEQUENCES: 15
;     CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Medlen & Carroll, LLP
;          STREET: 220 Montgomery Street, Suite 2200
;          CITY: San Francisco
;          STATE: California
;          COUNTRY: United States of America
;          ZIP: 94104
;     COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/244,752
;          FILING DATE:
;          CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: 08/735,609
;          FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;          NAME: Ingolia, Diane E.
```

```
;          REGISTRATION NUMBER: 40,027
;          REFERENCE/DOCKET NUMBER: UM-02484
;     TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (415) 705-8410
;          TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;          LENGTH: 35935 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: double
;          TOPOLOGY: linear
;     MOLECULE TYPE: other nucleic acid
;          DESCRIPTION: /desc = "DNA"
US-09-244-752-1

Query Match      8.3%; Score 15; DB 3; Length 35935;
Best Local Similarity 100.0%; Pred. No. 70;
Matches  15; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy     21 GACGCTGGAAGACGT 35
          ||||||||||||||
Db   6936 GACGCTGGAAGACGT 6950

RESULT 38
US-09-245-497-1
; Sequence 1, Application US/09245497
; Patent No. 6083750
; GENERAL INFORMATION:
;     APPLICANT: Chamberlain, Jeffrey S.
;     APPLICANT: Amalfitano, Andrea
;     APPLICANT: Hauser, Michael A.
;     APPLICANT: Kumar-Singh, Rajendra
;     APPLICANT: Hartigan-O'Connor, Dennis J.
;     TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;     NUMBER OF SEQUENCES: 15
;     CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Medlen & Carroll, LLP
;          STREET: 220 Montgomery Street, Suite 2200
;          CITY: San Francisco
;          STATE: California
;          COUNTRY: United States of America
;          ZIP: 94104
;     COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/245,497
;          FILING DATE:
;          CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: 08/735,609
;          APPLICATION NUMBER: <B> FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;          NAME: Ingolia, Diane E.
;          REGISTRATION NUMBER: 40,027
;          REFERENCE/DOCKET NUMBER: UM-02484
;     TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (415) 705-8410
;          TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;          LENGTH: 35935 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: double
;          TOPOLOGY: linear
;     MOLECULE TYPE: other nucleic acid
;          DESCRIPTION: /desc = "DNA"
US-09-245-497-1

Query Match      8.3%; Score 15; DB 3; Length 35935;
```

Best Local Similarity   100.0%; Pred. No. 70;
Matches 15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    21 GACGCTGGAAGACGT 35
         ||||||||||||||
Db  6936 GACGCTGGAAGACGT 6950
```

RESULT 39
US-09-409-670-43
; Sequence 43, Application US/09409670
; Patent No. 6113028
; GENERAL INFORMATION:
;   APPLICANT: IMLER, Jean-Luc
;   APPLICANT: MEHTALI, Majid
;   APPLICANT: PAVIRANI, Andrea
;   TITLE OF INVENTION: DEFECTIVE ADENOVIRUSES AND CORRESPONDING
;   TITLE OF INVENTION: COMPLEMENTATION LINES
;   NUMBER OF SEQUENCES: 43
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;   STREET: 1737 King Street, Suite 500
;   CITY: Alexandria
;   STATE: Virginia
;   COUNTRY: United States
;   ZIP: 22314-2756
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/409,670
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/379,452
;   FILING DATE: 26-JAN-1995
;   APPLICATION NUMBER: WO PCT/FR94/00624
;   FILING DATE: 27-MAY-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 93 06482
;   FILING DATE: 28-MAY-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Dadio, Susan M.
;   REGISTRATION NUMBER: 40,373
;   REFERENCE/DOCKET NUMBER: 029395-002
; INFORMATION FOR SEQ ID NO: 43:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 35935 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-09-409-670-43

Query Match   8.3%; Score 15; DB 3; Length 35935;
Best Local Similarity   100.0%; Pred. No. 70;
Matches 15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    21 GACGCTGGAAGACGT 35
         ||||||||||||||
Db  6936 GACGCTGGAAGACGT 6950
```

RESULT 40
US-09-562-919-1
; Sequence 1, Application US/09562919
; Patent No. 6451596
; GENERAL INFORMATION:
;   APPLICANT: Chamberlain, Jeffrey S.
;   Amalfitano, Andrea
;   Hauser, Michael A.
;   Kumar-Singh, Rajendra
;   Hartigan-O'Connor, Dennis J.
;   TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
;   NUMBER OF SEQUENCES: 15
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Medlen & Carroll, LLP
;   STREET: 220 Montgomery Street, Suite 2200
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: United States Of America
;   ZIP: 94104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/562,919
;   FILING DATE: 02-May-2000
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/735,609
;   FILING DATE: 23-Oct-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Ingolia, Diane E.
;   REGISTRATION NUMBER: 40,027
;   REFERENCE/DOCKET NUMBER: UM-02484
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 705-8410
;   TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 35935 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
;   DESCRIPTION: /desc = "DNA"
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-562-919-1

Query Match   8.3%; Score 15; DB 4; Length 35935;
Best Local Similarity   100.0%; Pred. No. 70;
Matches 15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    21 GACGCTGGAAGACGT 35
         ||||||||||||||
Db  6936 GACGCTGGAAGACGT 6950
```

RESULT 41
US-09-268-992-7
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
;   APPLICANT: Freimer, N.
;   APPLICANT: Chen, H.
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;   TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
;   FILE REFERENCE: 7853-138
;   CURRENT APPLICATION NUMBER: US/09/268,992
;   CURRENT FILING DATE: 1999-03-16
;   EARLIER APPLICATION NUMBER: 09/236,134
;   EARLIER FILING DATE: 1999-01-22
;   EARLIER APPLICATION NUMBER: 60/106,056
;   EARLIER FILING DATE: 1998-10-28
;   EARLIER APPLICATION NUMBER: 60/088,312
;   EARLIER FILING DATE: 1998-06-05
;   EARLIER APPLICATION NUMBER: 60/078,044
;   EARLIER FILING DATE: 1998-03-16
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7

```
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match   8.3%;  Score 15;  DB 4;  Length 72604;
Best Local Similarity 100.0%;  Pred. No. 70;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    161 TGATTGGAATGTTCT 175
          |||||||||||||||
Db  71962 TGATTGGAATGTTCT 71976
```

RESULT 42
US-09-657-474-7
```
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match   8.3%;  Score 15;  DB 4;  Length 72604;
Best Local Similarity 100.0%;  Pred. No. 70;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    161 TGATTGGAATGTTCT 175
          |||||||||||||||
Db  71962 TGATTGGAATGTTCT 71976
```

RESULT 43
US-08-584-040-7504/c
```
; Sequence 7504, Application US/08584040
; Patent No. 6346398
; GENERAL INFORMATION:
; APPLICANT: Pavco, Pamela
; APPLICANT: McSwiggen, James
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: TREATMENT OF DISEASES OR
; TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
; TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 8502
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/584,040
; FILING DATE: January 11, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,974
; FILING DATE: October 26, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 218/064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX:   (213) 955-0440
; TELEX:     67-3510
; INFORMATION FOR SEQ ID NO: 7504:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-584-040-7504

Query Match   7.7%;  Score 14;  DB 4;  Length 17;
Best Local Similarity 100.0%;  Pred. No. 2.1e+02;
Matches  14;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    34 GTCTTCCGAAGGAT 47
         ||||||||||||||
Db    17 GTCTTCCGAAGGAT 4
```

RESULT 44
US-08-584-040-7505/c
```
; Sequence 7505, Application US/08584040
; Patent No. 6346398
; GENERAL INFORMATION:
; APPLICANT: Pavco, Pamela
; APPLICANT: McSwiggen, James
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: TREATMENT OF DISEASES OR
; TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
; TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 8502
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
```

```
;        MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;        MEDIUM TYPE: storage
;        COMPUTER: IBM Compatible
;        OPERATING SYSTEM: IBM P.C. DOS 5.0
;        SOFTWARE: Word Perfect 5.1
;   CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/584,040
;        FILING DATE: January 11, 1996
;        CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: 60/005,974
;        FILING DATE: October 26, 1995
;   ATTORNEY/AGENT INFORMATION:
;        NAME: Warburg, Richard J.
;        REGISTRATION NUMBER: 32,327
;        REFERENCE/DOCKET NUMBER: 218/064
;   TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (213) 489-1600
;        TELEFAX: (213) 955-0440
;        TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 7505:
;        SEQUENCE CHARACTERISTICS:
;        LENGTH: 17 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
US-08-584-040-7505

Query Match     7.7%;  Score 14;  DB 4;  Length 17;
Best Local Similarity  100.0%;  Pred. No. 2.1e+02;
Matches  14;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       34 GTCTTCCGAAGGAT 47
            ||||||||||||||
Db       16 GTCTTCCGAAGGAT 3


RESULT 45
US-08-062-368-12
; Sequence 12, Application US/08062368
; Patent No. 5491086
;   GENERAL INFORMATION:
;        APPLICANT: Gelfand, David H.
;        APPLICANT: Wang, Alice M.
;        TITLE OF INVENTION: Purified Thermostable Nucleic Acid
;        TITLE OF INVENTION: Polymeraees Enzyme From Pyrodictium Species
;        NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Hoffmann-La Roche Inc.
;        STREET: 340 Kingsland Street
;        CITY: Nutley
;        STATE: New Jersey
;        COUNTRY: U.S.A.
;        ZIP: 07110
;   COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/062,368
;        FILING DATE: 19930514
;        CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;        NAME: Sias Ph.D., Stacey R.
;        REGISTRATION NUMBER: 32,630
;        REFERENCE/DOCKET NUMBER: 8584
;   TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (510) 814-2863
;        TELEFAX: (510) 814-2977
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;        LENGTH: 24 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
US-08-062-368-12

Query Match     7.7%;  Score 14;  DB 1;  Length 24;
Best Local Similarity  100.0%;  Pred. No. 2.2e+02;
Matches  14;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       87 AGCTGAGCAAGAGG 100
            ||||||||||||||
Db        4 AGCTGAGCAAGAGG 17


Search completed: June  9, 2003, 12:46:53
Job time : 98 secs
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 10:06:00 ; Search time 151 Seconds
                              (without alignments)
                              2699.415 Million cell updates/sec

Title:          US-09-550-163-1_COPY_80_260
Perfect score:  181
Sequence:       1 actttatccaattcacaca.........gattggaatgtccttca 181

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       2185239 seqs, 1125999159 residues

Word size :     0

Total number of hits satisfying chosen parameters:          4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :      N.Geneseq_101002:*
     1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
     2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
     3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
     4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
     5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
     6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
     7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
     8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
     9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
    10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
    11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
    12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
    13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
    14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
    15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
    16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
    17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
    18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
    19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
    20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
    21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
    22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
    23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
    24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	181	100.0	181	22	AAI24432	Probe #14365 for g
2	181	100.0	181	22	AAI09965	Probe #9956 used t
3	181	100.0	181	22	AAS00245	Human potassium ch
4	181	100.0	181	22	ABA44797	Human breast cell
5	181	100.0	181	22	ABA55252	Human foetal liver
6	181	100.0	181	22	ABA24997	Probe #3463 for ge
7	181	100.0	181	22	AAK03508	Human brain marrow
8	181	100.0	181	22	AAK28962	Human bone marrow
9	181	100.0	181	22	AAI13549	Probe #3482 for ge

No.	c	Length	%	Score	DB	ID	Description
10		181	100.0	450	22	AA134911	Probe #3597 used t
11		181	100.0	450	22	AA103438	Probe #3429 used t
12		181	100.0	450	22	ABS03496	Human genome-deriv
13		181	100.0	471	22	AA115256	Probe #5189 for ge
14		181	100.0	471	22	AA104990	Probe #4981 used t
15		181	100.0	600	22	AAF80269	Nucleotide sequenc
16		181	100.0	600	22	ABA09192	Human MiRP1 homolo
17		181	100.0	655	22	AAK52645	Human polynucleoti
18		181	100.0	732	22	AAK51661	Human polynucleoti
19		181	100.0	732	21	AAC64071	Human potassium ch
20		181	100.0	732	21	ABK86573	cDNA encoding huma
21		181	100.0	732	24	AAD35169	Human KCNE2 mutant
22		181	100.0	732	24	AAD35170	Human KCNE2 wild t
23		165	91.2	732	24	AAC64086	Human potassium ch
24		165	91.2	732	24	AAD35173	Human KCNE2 mutant
25		163	90.1	732	21	AAC64085	Human potassium ch
26		163	90.1	732	21	AAD35172	Human KCNE2 mutant
27		162	89.5	732	21	AAC64083	Human potassium ch
28		154	85.1	732	21	AAC64084	Human potassium ch
29		154	85.1	732	21	AAD35171	Human KCNE2 mutant
30		134	74.0	312	22	ABA49938	Human breast cell
31		134	74.0	312	22	ABA67856	Human foetal liver
32		134	74.0	312	22	ABA34913	Probe #13379 for g
33		134	74.0	312	22	AAK16264	Human brain expres
34		134	74.0	312	22	AAK42008	Human bone marrow
35		134	74.0	312	22	AAI22773	Probe #12706 for g
36		134	74.0	312	22	AAI48075	Probe #16761 used
37		134	74.0	312	22	AAI08446	Probe #8437 used t
38		134	74.0	312	24	ABS16039	Human genome-deriv
39	c	26	14.4	65	24	ABN29241	Rat spliced transc
40	c	26	14.4	372	22	AAS00246	Rat potassium chan
41	c	26	14.4	468	23	AAC64072	Rat potassium chan
42		23	12.7	21	21	AAC64079	PCR primer SEQ ID
43	c	21	11.6	21	21	AAC64078	PCR primer SEQ ID
44		21	11.6	21	24	AAD35176	Human KCNE2 gene a
45		20	11.0	20	21	AAF80272	Primer used to amp
46	c	20	11.0	20	24	AAD35175	Human KCNE2 gene a
47		18	9.9	688	23	AAS82550	DNA encoding novel
48		18	9.9	714	24	ABL07585	Human ovarian canc
49	c	18	9.9	743	23	AAS79957	DNA encoding novel
50	c	18	9.9	782	24	ABA06311	Soy bean SCN/SCS r
51		18	9.9	996	24	ABQ68798	Listeria monocytog
52	c	18	9.9	1841	17	AAT30651	P. gingivalis haem
53		18	9.9	1841	19	AAV58872	Haemagglutinin pro
54	c	18	9.9	2708	21	AAQ86784	DNA encoding Phosp
55		18	9.9	4985	24	ABO71005	Listeria monocytog
56		18	9.9	50000	23	ABL56201	AmEPV genome fragm
57	c	18	9.9	2944528	24	ABA03041	Listeria monocytog
58	c	17	9.4	243	21	AAC24301	Human secreted pro
59	c	17	9.4	331	21	AAF10129	Fusarium venenatum
60		17	9.4	374	21	ABL84041	Human ovarian canc
61		17	9.4	375	24	ABL84044	Human ovarian canc
62		17	9.4	396	21	ABN16910	Murine ORFX polynuc
63	c	17	9.4	2197	22	AAH47011	Murine mSPP1 prote
64		17	9.4	2384	22	AAH17956	Human cDNA sequenc
65	c	17	9.4	2693	23	AAH29840	S cerevisiae apopt
66	c	17	9.4	2708	16	AAQ86784	DNA encoding Phosp
67	c	17	9.4	4882	23	ABL02818	Drosophila melanog
68	c	17	9.4	8530	24	ABL33755	Human immune syste
69	c	17	9.4	8530	23	ABK31415	Signal transductio
70	c	17	9.4	9664	23	ABL17966	Drosophila melanog
71		17	9.4	32189	23	AAL04670	Human reproductive
72		17	9.4	32189	23	ABL97577	Human testicular a
73	c	17	9.4	75899	24	ABK85261	Human genomic DNA
74		17	9.4	465237	24	ABQ87681	Human oestrogen re
75		17	9.4	465237	24	ABA09193	Human oestrogen re
76		16	8.8	60	24	ABN46878	Human spliced tran
77		16	8.8	195	21	AA293816	Open reading frame
78		16	8.8	270	21	AAC14966	Human secreted pro
79		16	8.8	363	23	ABV06143	Human prostate exp
80	c	16	8.8	381	21	AAH30648	Human colon cancer
81		16	8.8	457	21	AAC38358	Zea mays DNA fragm
82		16	8.8	519	21	AA293818	Open reading frame

No.	c			Length	Accession	Description
83	c	16	8.8	545	AAH31082	Human colon cancer
84	c	16	8.8	622	AAZ80297	Human colon cancer
85		16	8.8	761	ABL19125	Drosophila melanog
86		16	8.8	795	AAK82908	Human immune/haema
87		16	8.8	937	AAH32014	Human olfactory re
88		16	8.8	1074	AAH32005	Human olfactory re
89	c	16	8.8	1100	AAC75819	Human ORFX ORF1374
90		16	8.8	1412	AAI58930	Human polynucleoti
91		16	8.8	1529	AAI60715	Human polynucleoti
92		16	8.8	1529	AAI60716	Human polynucleoti
93		16	8.8	1559	AAI58929	Human polynucleoti
94		16	8.8	1593	AAF29142	Polynucleotide seq
95		16	8.8	1665	AAI99104	Human excretory re
96		16	8.8	1665	AAI63454	Human kidney relat
97	c	16	8.8	1750	AAC47892	Arabidopsis thalia
98		16	8.8	1902	AAF58105	Arabidopsis thalia
99		16	8.8	1977	AAC43310	Arabidopsis thalia
100		16	8.8	2112	AAC43093	Arabidopsis thalia
101		16	8.8	2228	AAH16361	Human cDNA sequenc
102		16	8.8	2244	ABL65864	Lung cancer relate
103		16	8.8	2244	ABL66528	Lung cancer relate
104		16	8.8	2278	ABL18693	Drosophila melanog
105		16	8.8	2326	AAN30174	Sequence of the L1
106	c	16	8.8	2516	AAQ56832	Small round virus
107	c	16	8.8	2802	AAS69814	DNA encoding novel
108	c	16	8.8	2815	ABL19124	Drosophila melanog
109		16	8.8	2880	AAC78021	Human cancer assoc
110		16	8.8	2890	AAZ93781	Human desmin promo
111		16	8.8	2893	AAA38444	Human desmin gene
112		16	8.8	2900	ABK63943	cDNA encoding huma
113		16	8.8	2970	AAX20290	Borrelia burgdorfe
114	c	16	8.8	3316	AAC90486	Barley EDR1 cDNA.
115	c	16	8.8	3582	ABL26142	Drosophila melanog
116	c	16	8.8	4294	AAK86072	Human immune/haema
117		16	8.8	5004	ABL18692	Drosophila melanog
118		16	8.8	5807	ABL33155	Human immune syste
119	c	16	8.8	5962	ABL33286	Human immune syste
120		16	8.8	5993	AAK53105	Human polynucleoti
121	c	16	8.8	5993	AAK53106	Human polynucleoti
122	c	16	8.8	6047	ABL65188	Human cDNA differe
123	c	16	8.8	6047	ABK87063	Lung cancer relate
124	c	16	8.8	6567	AAK52121	Human polynucleoti
125	c	16	8.8	7170	ABL11029	Drosophila melanog
126	c	16	8.8	7722	AAQ12023	Norwalk virus geno
127		16	8.8	7753	AAQ56826	Norwalk virus stra
128		16	8.8	8283	ABL10086	Drosophila melanog
129	c	16	8.8	9826	ABL11028	Human polynucleoti
130	c	16	8.8	11990	ABL61824	Colon adenocarcino
131	c	16	8.8	11990	ABL69703	Prostate cancer re
132	c	16	8.8	14796	AAV27941	Survivin gene. Ho
133	c	16	8.8	14796	AAS21523	DNA encoding human
134	c	16	8.8	14796	AAH47531	Human Her-3 genomi
135		16	8.8	14796	ABN96924	Gene #3422 used to
136		16	8.8	14796	ABL65663	Lung cancer relate
137		16	8.8	14796	ABL66326	Lung cancer relate
138		16	8.8	14796	ABL68655	Kidney cancer rela
139	c	16	8.8	31314	AAL37510	Human musculoskele
140		16	8.8	32192	AAK89452	Human digestive sy
141		16	8.8	53585	AAX20251	Borrelia burgdorfe
142		16	8.8	144460	AAZ93815	Olfactory receptor
143	c	16	8.8	267156	ABL68560	Kidney cancer rela
144	c	16	8.8	640681	ABA92787	Buchnera sp. genom
145		16	8.8	2365589	24 ABA90521	Genomic sequence o
146		16	8.8	2944528	24 ABA03041	Listeria monocytog
147		15	8.3	20	20 AAZ02465	PCR primer used to
148		15	8.3	81	22 ABA72048	Human foetal liver
149		15	8.3	81	22 ABA38006	Probe #16472 for g
150		15	8.3	81	22 AAK20473	Human brain expres
151		15	8.3	81	22 AAK46614	Human bone marrow
152		15	8.3	81	22 AAI25681	Probe #1514 for g
153		15	8.3	81	22 AAI52450	Probe #21136 used
154		15	8.3	81	24 ABS20983	Human genome-deriv
155		15	8.3	158	24 AAI69095	Activated T-cell d

No.	c			Position	Len	Accession	Description
156	c	15	8.3	161	21	AAC28041	Human secreted pro
157	c	15	8.3	230	21	AAA41901	Human secreted exp
158	c	15	8.3	246	21	AAC93969	Cat flea hindgut a
160		15	8.3	271	22	ABA49762	Human breast cell
161		15	8.3	271	22	ABA67675	Human foetal liver
162		15	8.3	271	22	ABA34742	Probe #13208 for g
163		15	8.3	271	22	AAK16092	Human brain expres
164		15	8.3	271	22	AAK41834	Human bone marrow
165		15	8.3	271	22	AAI22600	Probe #12533 for g
166		15	8.3	271	22	AAI47892	Probe #16578 used
167		15	8.3	271	22	AAI08276	Probe #8267 used t
168	c	15	8.3	273	24	AAA81791	N. meningitidis pa
169		15	8.3	274	22	ABA48743	Human breast cell
170		15	8.3	274	22	ABA66653	Human foetal liver
171		15	8.3	274	22	ABA68846	Human foetal liver
172		15	8.3	274	22	ABA33718	Probe #12184 for g
173		15	8.3	274	22	ABA33914	Probe #12380 for g
174		15	8.3	274	22	AAK15082	Human brain expres
175		15	8.3	274	22	AAK15280	Human brain expres
176		15	8.3	274	22	AAK40810	Human bone marrow
177		15	8.3	274	22	AAK41004	Human bone marrow
178		15	8.3	274	22	AAI21580	Probe #11513 for g
179		15	8.3	274	22	AAI21774	Probe #11707 for g
180		15	8.3	274	22	AAI46869	Probe #15555 used
181		15	8.3	274	22	AAI47056	Probe #15742 used
182		15	8.3	274	22	AAI07274	Probe #7265 used t
183		15	8.3	274	22	AAI07459	Probe #7450 used t
184		15	8.3	274	24	ABS14764	Human genome-deriv
185		15	8.3	274	24	ABS14975	Human genome-deriv
186	c	15	8.3	295	24	ABL71174	Corn tassel-derive
187		15	8.3	328	21	AAA43002	Human secreted exp
188	c	15	8.3	346	24	ABL63251	Breast cancer rela
189		15	8.3	369	22	AAS38319	Novel human diagno
190		15	8.3	378	24	ABK64666	Human benign prost
191		15	8.3	389	24	ABN63064	Human cancer relat
192		15	8.3	402	24	ABQ91009	M. capsulatus gene
193		15	8.3	404	23	ABV08094	Human prostate exp
194		15	8.3	407	20	AAK53916	Murine translation
195		15	8.3	412	20	AAV87032	EST clone BG173.
196	c	15	8.3	415	24	ABN25363	Human ORFX polynuc
197		15	8.3	420	19	AAV10365	FRAP homolog. Cry
198		15	8.3	425	20	AAX41311	Human secreted pro
199	c	15	8.3	429	24	ABN26114	Human ORFX polynuc
200		15	8.3	433	23	ABV00919	Human prostate exp
201	c	15	8.3	453	22	AAL07207	Human reproductive
202		15	8.3	455	22	ABV98010	Human prostate exp
203		15	8.3	455	22	ABA43636	Human breast cell
204		15	8.3	455	22	ABA54085	Human foetal liver
205		15	8.3	455	22	ABA23837	Probe #203 for ge
206		15	8.3	455	22	AAK02360	Human brain expres
207		15	8.3	455	22	AAK27802	Human bone marrow
208		15	8.3	455	22	AAI12387	Probe #2320 for ge
209		15	8.3	455	22	AAI33741	Probe #2427 used t
210		15	8.3	455	22	AAI02299	Probe #2290 used t
211		15	8.3	465	24	ABS02256	Human genome-deriv
212		15	8.3	472	23	ABV98010	Human prostate exp
213		15	8.3	472	22	ABA44615	Human breast cell
214		15	8.3	472	22	ABA55067	Human foetal liver
215		15	8.3	472	22	ABA24821	Probe #3287 for ge
216		15	8.3	472	22	AAK03332	Human brain expres
217		15	8.3	472	22	AAK28783	Human bone marrow
218		15	8.3	472	22	AAI13369	Probe #3302 for ge
219		15	8.3	472	22	AAI34722	Probe #3408 used t
220		15	8.3	472	22	AAI03263	Probe #3254 used t
221		15	8.3	475	24	ABS03308	Human genome-deriv
222		15	8.3	475	22	ABA54284	Human foetal liver
223		15	8.3	475	22	ABA24040	Probe #2506 for ge
224		15	8.3	475	22	AAK28003	Human brain expres
225		15	8.3	475	22	AAI12586	Human bone marrow
226		15	8.3	475	22	AAI13933	Probe #2519 for ge
227		15	8.3	475	22	AAI02491	Probe #2619 used t
228		15	8.3	475	24	ABS02474	Probe #2482 used t / Human genome-deriv

c	#			Accession	len	pos	Description
c	229	8.3	15	AAS90514	23	483	DNA encoding novel
	230	8.3	15	AAL23311	22	498	Human breast cance
	231	8.3	15	AAH65529	22	516	C glutamicum codin
	232	8.3	15	AAH11971	22	523	Human cDNA clone (
	233	8.3	15	ABK53522	24	535	Human eosinophil-m
	234	8.3	15	ABA62924	22	541	Human foetal liver
	235	8.3	15	ABA29852	22	541	Probe #8318 for ge
	236	8.3	15	AAK10877	22	541	Human brain expres
	237	8.3	15	AAK36738	22	541	Human bone marrow
	238	8.3	15	AAI17589	22	541	Probe #7522 for ge
	239	8.3	15	ABS10736	22	541	Probe #11191 used
	240	8.3	15	ABA63153	24	555	Human genome-deriv
	241	8.3	15	ABA30402	22	555	Human foetal liver
	242	8.3	15	AAK11590	22	555	Probe #8868 for ge
	243	8.3	15	AAK37358	22	555	Human brain expres
	244	8.3	15	AAI18183	22	555	Human bone marrow
	245	8.3	15	AAI43202	22	555	Probe #8116 for ge
	246	8.3	15	ABS11346	24	555	Probe #11888 used
	247	8.3	15	AAH11603	22	559	Human genome-deriv
	248	8.3	15	AAF21624	21	563	Human cDNA clone (
	249	8.3	15	ABA59505	22	563	Human breast and o
	250	8.3	15	ABA28126	22	593	Human foetal liver
	251	8.3	15	AAK07777	22	593	Probe #6592 for ge
	252	8.3	15	AAK33644	22	593	Human brain expres
	253	8.3	15	AAI16484	22	593	Human bone marrow
	254	8.3	15	AAI39367	22	593	Probe #6417 for ge
	255	8.3	15	ABS08548	24	593	Probe #8053 used t
	256	8.3	15	AAC10902	21	593	Human genome-deriv
	257	8.3	15	ABV52042	23	618	Human secreted pro
c	258	8.3	15	ABQ68568	24	623	Human prostate exp
	259	8.3	15	AAF12143	21	631	Listeria monocytog
c	260	8.3	15	AAH03550	22	673	Aspergillus oryzae
c	261	8.3	15	AAF14273	21	715	Human cDNA clone (
	262	8.3	15	AAI94730	22	734	Aspergillus oryzae
	263	8.3	15	AAC10881	21	744	Human neuroblastom
	264	8.3	15	ABL24211	23	745	Drosophila melanog
	265	8.3	15	AAI97410	22	780	Human neuroblastom
	266	8.3	15	ABN20554	24	800	Human ORFX polynuc
c	267	8.3	15	ABK54093	24	801	DNA encoding Neiss
	268	8.3	15	AAS62134	24	810	Porcine muscular s
	269	8.3	15	AAA50074	21	813	Neisseria meningit
c	270	8.3	15	AAF22753	22	816	Human gastric canc
	271	8.3	15	AAS61966	24	847	Porcine muscular s
c	272	8.3	15	AAH06653	22	853	Human cDNA clone (
	273	8.3	15	AAC42611	21	860	Arabidopsis thalia
	274	8.3	15	ABQ91776	24	881	M. capsulatus gene
c	275	8.3	15	AAI88942	22	903	Human polynucleoti
	276	8.3	15	AAH53854	22	940	S. epidermidis ope
	277	8.3	15	AAX97761	20	942	Drosophila melanog
	278	8.3	15	AAF22751	22	959	Extended human sec
c	279	8.3	15	AAF91384	21	960	Human gastric canc
	280	8.3	15	AAF37764	24	1000	N. meningitidis (s
c	281	8.3	15	AAF22749	22	1005	DNA sequence upstr
	282	8.3	15	ABL04027	23	1005	Human gastric canc
c	283	8.3	15	ABK35625	24	1124	Drosophila melanog
	284	8.3	15	AAF27731	22	1135	cDNA sequence #16
c	285	8.3	15	AAC45543	21	1139	Human transport pr
	286	8.3	15	AAC40234	24	1164	Arabidopsis thalia
	287	8.3	15	AAC45887	21	1182	Arabidopsis thalia
	288	8.3	15	ABK35807	24	1225	cDNA sequence #198
	289	8.3	15	AAAS9060	21	1240	Extended human sec
	290	8.3	15	ABA94272	22	1240	Nucleotide sequenc
	291	8.3	15	AAC39269	21	1242	Adenovirus 5 tripa
	292	8.3	15	AAK98900	24	1242	DNA of Pantoea cit
	293	8.3	15	AAB91874	24	1251	Pantoea citrea per
	294	8.3	15	AAF16249	21	1256	Human protein enco
	295	8.3	15	AAH99650	22	1256	Human prostate can
c	296	8.3	15	AAC47451	21	1308	Arabidopsis thalia
c	297	8.3	15	AAC39269	21	1312	Arabidopsis thalia
	298	8.3	15	AAD05457	22	1346	Human secreted pro
	299	8.3	15	AAQ68861	24	1361	Listeria monocytog
	300	8.3	15	AAS26041	22	1396	Human cDNA encodin
c	301	8.3	15	AAD02200	22	1399	Chondrus crispus s
	302	8.3	15	AAS02046	22	1421	DNA encoding molec
c	303	8.3	15	AAF21615	21	1425	Human breast and o
	304	8.3	15	AAI60558	22	1465	Human polynucleoti
	305	8.3	15	AAI58772	22	1467	Human polynucleoti
c	306	8.3	15	AAV42998	19	1471	Streptococcus pneu
	307	8.3	15	ABK10185	22	1500	2,5-diketo-D-gluco
	308	8.3	15	AAF61079	22	1515	P. putida KT2440-a
	309	8.3	15	AAC49642	21	1558	Arabidopsis thalia
c	310	8.3	15	AAC46629	21	1703	Zea mays DNA fragm
	311	8.3	15	AAP26562	22	1728	DNA encoding human
	312	8.3	15	AAS53828	21	1733	Sequence encoding
c	313	8.3	15	AAA53146	21	1790	Arabidopsis thalia
c	314	8.3	15	AAH17051	22	1793	Human cDNA sequenc
	315	8.3	15	ABL29301	23	1820	Drosophila melanog
	316	8.3	15	ABN91712	21	1854	Staphylococcus epi
	317	8.3	15	AAH44629	22	1932	Human cathepsin 18
c	318	8.3	15	ABK34805	22	2059	Human cDNA for nov
c	319	8.3	15	AAV52307	19	2069	Streptococcus pneu
	320	8.3	15	ABL91309	24	2088	Chlamydia pneumoni
	321	8.3	15	AAH90025	22	2102	Human bone marrow
	322	8.3	15	ABV25190	23	2136	Human prostate exp
c	323	8.3	15	AAN91238	10	2220	B1 antigen. Toxop
c	324	8.3	15	ABL24210	23	2257	Drosophila melanog
c	325	8.3	15	AAH89912	21	2270	Human bone marrow
c	326	8.3	15	ABL19083	23	2389	Drosophila melanog
c	327	8.3	15	AAH44517	20	2435	Human reticulin 80
	328	8.3	15	AAZ27510	19	2446	DS9 gene sequence.
c	329	8.3	15	AAA71390	21	2446	Tobacco DS9 cDNA t
	330	8.3	15	AAF86349	22	2517	Murine cDNA encodi
c	331	8.3	15	AAS51796	23	2535	Staphylococcus aur
c	332	8.3	15	AAN70548	8	2565	2.6 kbp fragment o
c	333	8.3	15	AAQ24233	13	2607	Varicella-zoster v
c	334	8.3	15	AAS45431	21	2610	Staphylococcus aur
	335	8.3	15	AAD05409	22	2648	Human secreted pro
	336	8.3	15	ABN96803	24	2682	Gene #3301 used to
	337	8.3	15	ABL25117	23	2705	Drosophila melanog
c	338	8.3	15	AAA70255	21	2715	Plasmodium falcipa
	339	8.3	15	AAH14645	22	2747	Human cDNA sequenc
	340	8.3	15	AAH77272	21	2765	Arabidopsis indehi
	341	8.3	15	ABQ54870	24	2798	Human ovarian anti
	342	8.3	15	AAC59995	23	2799	Human secreted pro
	343	8.3	15	ABV21696	21	2854	Human prostate exp
	344	8.3	15	ABV27517	23	2854	Human prostate exp
c	345	8.3	15	AAH16259	22	2874	Human cDNA sequenc
	346	8.3	15	AAH16493	22	2897	Human cDNA sequenc
	347	8.3	15	ABL25108	23	2959	Drosophila melanog
	348	8.3	15	AAH73602	23	2979	Human WT1-interact
	349	8.3	15	ABL20076	22	3023	Drosophila melanog
	350	8.3	15	AAH17871	22	3102	Human cDNA sequenc
	351	8.3	15	AAH14656	22	3172	Human cDNA sequenc
	352	8.3	15	AAV26030	19	3258	Human transient re
	353	8.3	15	ABL04026	23	3321	Drosophila melanog
	354	8.3	15	AAS25917	21	3353	Human cDNA encodin
	355	8.3	15	AAS39217	22	3357	Type IV adenylyl c
	356	8.3	15	ABL21526	23	3661	Drosophila melanog
	357	8.3	15	ABL05274	23	3662	Drosophila melanog
	358	8.3	15	AAH77271	22	3856	Arabidopsis indehi
c	359	8.3	15	AAH54945	22	3983	S. epidermidis gen
	360	8.3	15	AAH54144	21	3990	S. epidermidis gen
	361	8.3	15	AAV74534	18	3997	Staphylococcus aur
	362	8.3	15	ABK84306	24	4064	Human cDNA differe
	363	8.3	15	ABL23594	23	4082	Drosophila melanog
	364	8.3	15	ABV21780	23	4085	Human prostate exp
	365	8.3	15	ABV27603	23	4085	Human prostate exp
	366	8.3	15	AAC77490	21	4211	Human ORFX ORF3045
	367	8.3	15	AAH98501	22	4341	Human EST-derived
c	368	8.3	15	AAS06386	22	4454	Vector pGX104 DNA
c	369	8.3	15	ABV25119	23	4531	Human prostate exp
c	370	8.3	15	AAA70129	21	4551	Plasmodium falcipa
	371	8.3	15	ABQ70997	24	4556	Listeria monocytog
	372	8.3	15	ABL04520	23	4637	Drosophila melanog
c	373	8.3	15	AAS94850	24	4760	Human DNA sequence
c	374	8.3	15	ABL25116	23	4771	Drosophila melanog

Flag	#			Pos	Len	Accession	Description
c	375	15	8.3	4927	24	ABQ70968	Listeria monocytog
c	376	15	8.3	4965	19	AAV52276	Streptococcus pneu
	377	15	8.3	5248	23	ABL10374	Drosophila melanog
c	378	15	8.3	5261	24	ABL70577	Chemically treated
c	379	15	8.3	5261	24	AAS61392	Human gene regulat
	380	15	8.3	5328	23	ABL32561	Human immune syste
	381	15	8.3	5557	23	ABL04036	Drosophila melanog
	382	15	8.3	5736	15	AAS87512	DNA encoding novel
	383	15	8.3	5774	22	AAS45485	Chemically pretrea
c	384	15	8.3	5774	24	ABK28411	DNA transcription
	385	15	8.3	5799	23	AAS78436	DNA encoding novel
	386	15	8.3	5799	23	AAS79370	Chemically pretrea
c	387	15	8.3	6022	22	AAS45375	Tumour suppressor
c	388	15	8.3	6022	22	AAS46398	Human immune syste
c	389	15	8.3	6022	24	ABL33073	DNA transcription
	390	15	8.3	6022	22	ABK28212	Human immune syste
c	391	15	8.3	6103	23	ABL34171	Signal transductio
	392	15	8.3	6103	24	ABK31513	Human immune syste
	393	15	8.3	6161	24	ABL32623	Drosophila melanog
c	394	15	8.3	6224	23	ABL28276	Human ORFX ORF2747
	395	15	8.3	6538	21	AAC77192	DNA encoding novel
	396	15	8.3	6637	23	AAS71042	Human immune syste
c	397	15	8.3	6881	24	ABL33380	DNA transcription
	398	15	8.3	6881	23	ABK28249	Drosophila melanog
	399	15	8.3	6883	23	ABL26736	Nucleotide sequenc
	400	15	8.3	7231	21	AAA59090	Nucleotide sequenc
	401	15	8.3	7231	15	ABA94286	Human chemically m
	402	15	8.3	7406	24	ABN80156	Human chemically t
	403	15	8.3	7434	24	AAD28387	Drosophila melanog
	404	15	8.3	7443	22	ABL06686	Human musculoskele
c	405	15	8.3	7601	22	AAL36533	Human musculoskele
c	406	15	8.3	7601	21	AAS09072	Nucleotide sequenc
	407	15	8.3	7960	22	ABA94274	Nucleotide sequenc
	408	15	8.3	7960	21	AAS09075	Nucleotide sequenc
	409	15	8.3	7989	23	ABA94277	Drosophila melanog
	410	15	8.3	7989	21	ABL02574	Nucleotide sequenc
	411	15	8.3	8043	21	AAS09071	Nucleotide sequenc
	412	15	8.3	8383	23	ABA94273	Nucleotide sequenc
	413	15	8.3	8383	23	AAS09091	Drosophila melanog
c	414	15	8.3	8484	22	ABL19082	Drosophila melanog
	415	15	8.3	9431	23	ABL07586	Human immune/haema
	416	15	8.3	9721	23	ABL21171	Nucleotide sequenc
c	417	15	8.3	9953	23	AAK79779	Drosophila melanog
c	418	15	8.3	10312	20	AAV33921	Human reproductive
	419	15	8.3	10332	23	ABL02358	Human testicular a
	420	15	8.3	10868	23	AAL05813	Human musculoskele
	421	15	8.3	11960	22	ABL98377	Human reproductive
	422	15	8.3	11960	22	AAL36567	Polynucleotide seq
	423	15	8.3	11991	22	AAL07205	DNA encoding novel
	424	15	8.3	11991	22	AAX20547	Drosophila melanog
	425	15	8.3	12027	23	AAS87478	Propionibacterium
	426	15	8.3	12260	22	ABL08874	Drosophila melanog
c	427	15	8.3	12292	23	AAS59537	Drosophila melanog
	428	15	8.3	12550	23	ABL29300	Streptococcus pneu
	429	15	8.3	12629	23	ABL21170	Drosophila melanog
	430	15	8.3	12666	19	AAV52270	Tumour suppressor
c	431	15	8.3	13133	23	ABL05324	Human gene regulat
c	432	15	8.3	13317	22	AAS46682	Polynucleotide seq
c	433	15	8.3	13317	24	AAS61367	Streptococcus pneu
c	434	15	8.3	13330	22	AAS87478	Human genomic DNA
c	435	15	8.3	13440	19	AAV52261	DNA encoding enzym
c	436	15	8.3	14176	22	AAS26670	Nucleotide sequenc
c	437	15	8.3	14245	21	AAZ45259	Mouse bone morphog
	438	15	8.3	14245	23	AAZ30356	Human immune/haema
	439	15	8.3	15144	18	AAT47715	Human immune/haema
	440	15	8.3	16273	22	AAK79625	Drosophila melanog
	441	15	8.3	16273	23	ABL06068	DNA encoding novel
c	442	15	8.3	16463	23	AAK80688	Tumour suppressor
c	443	15	8.3	17727	22	AAS87482	DNA encoding novel
c	444	15	8.3	19131	22	AAS46717	Human cardiovascul
c	445	15	8.3	19861	23	AAS87516	Drosophila melanog
c	446	15	8.3	21129	23	AAS36234	
c	447	15	8.3	21747	23	ABL09060	

Flag	#			Pos	Len	Accession	Description
	448	15	8.3	23071	22	AAS26699	Human genomic DNA
	449	15	8.3	23532	21	AAA81455	N. meningitidis pa
c	450	15	8.3	26410	24	AAK70623	Human immune/haema
c	451	15	8.3	30365	24	ABK49011	Adenoviral vector
	452	15	8.3	31183	22	AAD03963	Adenovirus Ad-5 de
	453	15	8.3	31446	21	AAC09088	AdPB-beta-galactos
	454	15	8.3	31672	18	ABK49010	Adenoviral vector
c	455	15	8.3	32026	24	AAT60559	Recombinant adenov
	456	15	8.3	32165	21	AAA09092	AdMMTV-beta-galact
	457	15	8.3	32165	21	AAA14723	Nucleotide sequenc
	458	15	8.3	32166	21	AAA09090	AdPSA-beta-galacto
	459	15	8.3	32166	22	AAC89170	AdRSVpHYDE region
	460	15	8.3	32167	21	AAC89170	Nucleotide sequenc
	461	15	8.3	32167	21	AAI14803	Partial sequence o
	462	15	8.3	32480	21	AA293332	Nucleotide sequenc
	463	15	8.3	32480	21	AAA59055	Adenovirus related
	464	15	8.3	32798	24	ABA94267	Replication defici
	465	15	8.3	32886	21	ABA97684	AdRSV-beta-galacto
	466	15	8.3	33023	24	AAC09086	Gene #351 used to
	467	15	8.3	33140	22	ABN96853	Genomic fragment #
	468	15	8.3	33592	22	AAF28536	Adenovirus anti-ca
	469	15	8.3	33669	22	AACB5018	Adenovirus anti-ca
	470	15	8.3	33988	22	AACB5020	Adenovirus anti-ca
	471	15	8.3	34302	21	AACB5023	Adenovirus vector
	472	15	8.3	34303	19	AAV07261	Adenoviral vector
	473	15	8.3	34341	22	AACB5019	Adenovirus anti-ca
	474	15	8.3	34382	20	AAX15627	Recombinant adenov
	475	15	8.3	34447	20	AAX07371	Adenovirus vector
	476	15	8.3	34448	22	AAC85021	Adenovirus anti-ca
c	477	15	8.3	34616	24	AAC49009	Adenoviral vector
	478	15	8.3	34737	22	AACB5024	Adenovirus anti-ca
	479	15	8.3	35000	18	AAT60557	Recombinant cis-ac
	480	15	8.3	35724	22	AAC85022	Recombinant adenov
	481	15	8.3	35871	24	AAD27972	Adenovirus anti-ca
	482	15	8.3	35935	19	AAV07258	Recombinant adenov
	483	15	8.3	35935	22	AACB5026	Adenovirus 5 genom
	484	15	8.3	35978	24	AAD27971	Complete nucleotid
	485	15	8.3	36114	24	AAC85025	Recombinant adenov
	486	15	8.3	36154	24	AAK51108	Adenovirus anti-ca
	487	15	8.3	36154	24	ABK51116	(+)lox(+)pol helpe
	488	15	8.3	36335	15	AAQ68003	DNA sequence of de
	489	15	8.3	36538	18	AAT60558	Ad2/-ORF6/PGK-CFTR
	490	15	8.3	37339	22	AAS15612	Recombinant trans-
	491	15	8.3	37391	22	ABK51122	Mouse osteocalcin
	492	15	8.3	38042	24	ABK51120	Delta HIX#3 vector
	493	15	8.3	40308	22	AAK68184	pD1940 vector DNA
	494	15	8.3	43671	23	ABL16160	Human immune/haema
c	495	15	8.3	44608	21	AAAB1495	Drosophila melanog
	496	15	8.3	45186	22	AAF60478	N. meningitidis pa
	497	15	8.3	47475	21	AAB1465	Wild-type human CT
	498	15	8.3	51402	22	AAK72363	N. meningitidis pa
	499	15	8.3	72604	20	AA210752	Human immune/haema
	500	15	8.3	72604	24	AAH43231	Genomic sequence o
	501	15	8.3	92934	21	AAAB1473	Human HKNG1 gene.
	502	15	8.3	110608	24	ABK83572	N. meningitidis pa
	503	15	8.3	124884	22	AAH74201	Human cDNA differe
c	504	15	8.3	125157	22	AAH72202	Nucleotide sequenc
c	505	15	8.3	154465	24	AAD28763	Nucleotide sequenc
c	506	15	8.3	158245	24	AAD28762	Human AKAP allelic
c	507	15	8.3	161425	22	AAH02340	Human AKAP allelic
c	508	15	8.3	162025	22	AAH02339	Human AKAP10 gene
c	509	15	8.3	162025	24	AAD28758	Human AKAP10 gene
c	510	15	8.3	162025	24	AAD28759	Human AKAP allelic
c	511	15	8.3	172325	21	AAF21613	Human AKAP allelic
	512	15	8.3	273254	21	AAC81914	Neisseria meningit
	513	15	8.3	349980	21	AAF21607	Chlamydia pneumoni
	514	15	8.3	349980	21	AAF21608	Neisseria meningit
	515	15	8.3	349980	21	AAF21610	Neisseria meningit
	516	15	8.3	349980	21	AAF21612	Neisseria meningit
	517	15	8.3	349980	22	AAH41226	Neisseria meningit
	518	15	8.3	349980	22	AAH68525	Pyrococcus abyssi
	519	15	8.3	368004	24	ABL57909	C glutamicum Codin
c	520	15	8.3				Human transporter

us-09-550-163-1_copy_80_260.oligo.rng

	Seq		Score	Pos	Accession	Len	Description
	521	15	8.3	837096	AAA81489	21	N. meningitidis pa
	522	15	8.3	1038602	AAA201425	20	Complete genome se
c	523	15	8.3	1038602	AAA201425	20	Complete genome se
c	524	15	8.3	1230025	AAX91990	20	Nucleotide sequenc
c	525	15	8.3	1437668	AAA81490	21	N. meningitidis B
	526	14	7.7	17	AAX74754	18	Mouse flt-1 VEGF r
	527	14	7.7	17	AAX74755	18	Mouse flt-1 VEGF r
c	528	14	7.7	19	AAT71821	21	Human biallelic ma
	529	14	7.7	20	AAQ36627	14	Truncated hKL 5' p
	530	14	7.7	21	AAQ37901	14	Beta-casein sequen
c	531	14	7.7	21	AAA52786	21	Streptococcus pneu
	532	14	7.7	21	AAC64081	21	PCR primer SEQ ID
	533	14	7.7	22	AAQ73829	14	P. occultum DNA po
c	534	14	7.7	24	AAH49081	22	Human LDLR gene as
	535	14	7.7	29	AAF56213	22	DNA binding protei
	536	14	7.7	40	AAT87225	24	Interleukin-4 2'NH
c	537	14	7.7	41	ABK48554	18	Forward PCR primer
	538	14	7.7	50	AAL34586	22	Human SNP oligonuc
	539	14	7.7	60	ABN37524	24	Human spliced tran
	540	14	7.7	65	ABN29078	24	Rat spliced transc
c	541	14	7.7	65	ABN30171	24	Rat spliced transc
c	542	14	7.7	65	ABN52632	24	Mouse spliced tran
	543	14	7.7	71	AAQ76344	15	Immunogenic protei
	544	14	7.7	71	AAQ76345	15	Immunogenic protei
c	545	14	7.7	87	AAV75424	18	Staphylococcus aur
c	546	14	7.7	90	AAC83177	22	PCR primer used in
c	547	14	7.7	90	AAC83178	22	PCR primer used in
c	548	14	7.7	96	AAC83181	22	PCR primer used in
	549	14	7.7	98	AAI65431	22	Nucleotide sequenc
	550	14	7.7	101	AAT58970	18	Bacterial expressi
	551	14	7.7	101	AAT58969	18	Bacterial expressi
	552	14	7.7	102	AAF15083	21	Trichoderma reesei
c	553	14	7.7	102	AAA45576	21	Human secreted exp
	554	14	7.7	117	AAH85705	20	Human single nucle
c	555	14	7.7	120	ABL22361	23	Drosophila melanog
	556	14	7.7	133	ABA48806	22	Human breast cell
	557	14	7.7	133	ABA66723	22	Human foetal liver
	558	14	7.7	133	ABA33781	22	Probe #12247 for g
	559	14	7.7	133	AAK15148	21	Human breast cance
	560	14	7.7	133	AAK40877	22	Human bone marrow
	561	14	7.7	133	AAI21642	22	Probe #11575 for g
	562	14	7.7	133	AAI46931	22	Probe #15617 used
c	563	14	7.7	133	AAI07334	22	Probe #7325 used t
	564	14	7.7	133	ABS14834	24	Human genome-deriv
c	565	14	7.7	139	AAC21997	21	Human secreted pro
	566	14	7.7	143	AAQ33920	13	Sequence upstream
	567	14	7.7	162	ABA51695	24	Portion of the pQE
c	568	14	7.7	165	AAA45244	21	Human secreted exp
c	569	14	7.7	171	ABA50558	22	Human breast cell
	570	14	7.7	171	AAB68516	22	Human foetal liver
c	571	14	7.7	171	ABA35497	22	Probe #13963 for g
	572	14	7.7	171	AAK16884	22	Human brain expres
	573	14	7.7	171	AAK42654	22	Human bone marrow
	574	14	7.7	171	AAI23408	22	Probe #13341 for g
	575	14	7.7	171	AAI49728	22	Probe #17414 used
	576	14	7.7	171	AAI09035	20	Probe #9026 used t
c	577	14	7.7	171	ABS16706	24	Human genome-deriv
	578	14	7.7	179	AAZ32175	21	Human low density
	579	14	7.7	182	AAC14138	21	Human secreted pro
	580	14	7.7	186	ABA49930	22	Human breast cell
	581	14	7.7	186	ABA67848	22	Human foetal liver
	582	14	7.7	186	ABA34907	22	Probe #13373 for g
	583	14	7.7	186	AAK16258	22	Human brain expres
	584	14	7.7	186	AAI22766	22	Probe #12699 for g
	585	14	7.7	186	AAI48068	22	Probe #16754 used
	586	14	7.7	186	AAI04318	20	Probe #8430 used t
	587	14	7.7	186	ABS16031	24	Human genome-deriv
	588	14	7.7	195	ABL63245	24	Breast cancer rela
	589	14	7.7	197	AAT23269	16	Human gene signatu
	590	14	7.7	197	ABV61536	23	Human prostate exp
c	591	14	7.7	202	AAC15201	21	Human secreted pro
	592	14	7.7	207	AAN91067	10	Control sequence N
	593	14	7.7	208	AAT23257	16	Human gene signatu

Description	Accession	Len	Seq	Score			
Prostate specific	AAT90074	18	246	7.7	14	c	594
Plasmid p16074a ex	AA257923	21	247	7.7	14		595
Repressor plasmid	AAV02345	18	250	7.7	14	c	596
Human biallelic po	AAX11288	19	251	7.7	14		597
Human biallelic po	AAX11289	19	251	7.7	14		598
Human biallelic po	AAX11290	19	251	7.7	14		599
Human ovarian canc	ABL87458	24	253	7.7	14	c	600
Human ORF2476 cDNA	ABN77529	24	256	7.7	14		601
Drosophila melanog	ABL18965	23	264	7.7	14		602
Staphylococcus aur	AAS50735	23	266	7.7	14	c	603
Mycobacterium tube	ABQ62567	22	277	7.7	14		604
Human prostate exp	ABV08458	23	277	7.7	14	c	605
Human secreted pro	AAC03683	21	286	7.7	14		606
Human musculoskele	AAL35440	22	289	7.7	14	c	607
Corn tassel-derive	ABL71172	24	292	7.7	14		608
Human ORFX polynuc	ABN18198	24	298	7.7	14	c	609
Human gene express	AAZ14092	20	300	7.7	14		610
Human cancer cell	AAX98529	20	300	7.7	14		611
Corn tassel-derive	ABL71253	24	300	7.7	14		612
Human nervous syst	ABA21459	22	307	7.7	14		613
Human secreted pro	AAC04417	21	312	7.7	14	c	614
Human prostate exp	ABV17487	23	312	7.7	14		615
DNA encoding novel	AAS83502	22	321	7.7	14	c	616
Human glycoprotein	ABN78963	24	324	7.7	14		617
Human immune/haema	AAK63542	22	325	7.7	14		618
Arabidopsis thalia	ABQ86123	24	329	7.7	14		619
Human bone marrow	AAK49369	24	330	7.7	14		620
Human stem cell an	AAV38051	19	335	7.7	14		621
Human nervous syst	ABA16441	22	337	7.7	14	c	622
Human nervous syst	ABA16442	22	337	7.7	14	c	623
Human nervous syst	ABA16443	22	337	7.7	14	c	624
Human nervous syst	ABA16444	22	337	7.7	14	c	625
Human nervous syst	ABA11903	22	338	7.7	14	c	626
Human olfactory re	AAH32182	22	342	7.7	14		627
Human single nucle	AAH87096	20	343	7.7	14		628
Human single nucle	AAH87385	24	343	7.7	14		629
Human ORFX polynuc	ABN17479	24	343	7.7	14	c	630
Human breast cance	AAI89633	22	344	7.7	14		631
Human breast cance	AAL25761	22	345	7.7	14		632
Human prostate exp	ABV34239	23	347	7.7	14		633
Human prostate exp	ABV49863	23	350	7.7	14		634
Novel human polyns	AAF15466	22	350	7.7	14		635
Human breast cance	AAL25877	22	351	7.7	14	c	636
Human breast cance	AAL08132	22	354	7.7	14	c	637
Human breast cance	AAL19042	22	355	7.7	14	c	638
Human polynucleoti	AAI91593	22	357	7.7	14	c	639
Human breast cance	AAL08015	22	358	7.7	14		640
Human breast cance	AAL18926	22	358	7.7	14	c	641
Human polynucleoti	AAI82878	22	358	7.7	14		642
Human secreted pro	AAC30528	21	359	7.7	14		643
Lung carcinoma CDN	AAP33525	24	364	7.7	14	c	644
Human cancer relat	ABN62908	24	366	7.7	14		645
Human contig polyn	AAS44823	22	368	7.7	14		646
Probe #14365 for g	AAI24432	22	368	7.7	14	c	647
Probe #9956 used t	AAI09965	14	372	7.7	14	c	648
Human potassium ch	AAS00245	22	372	7.7	14		649
Human ORFX polynuc	ABN19308	24	372	7.7	14	c	650
Human prostate exp	ABV17835	23	374	7.7	14	c	651
Human secreted pro	AAC11776	21	375	7.7	14		652
Human breast cance	AAL08236	22	376	7.7	14	c	653
DNA encoding anti-	AAS03450	22	378	7.7	14		654
Human ORFX polynuc	ABN26451	24	378	7.7	14	c	655
Human prostate exp	ABV38363	23	380	7.7	14		656
Human prostate exp	ABV42495	23	387	7.7	14	c	657
Human prostate exp	ABV42586	23	387	7.7	14	c	658
Gene #816 used to	ABN94318	14	391	7.7	14		659
cDNA encoding rat	AAZ61631	22	393	7.7	14	c	660
Skin cell cDNA, SE	AAC99564	21	393	7.7	14	c	661
Rat cDNA isolated	ABLJ34716	22	393	7.7	14	c	662
Human nervous syst	ABA21093	22	396	7.7	14		663
Human prostate exp	ABV08312	23	396	7.7	14	c	664
Human secreted pro	ABV06769	23	398	7.7	14	c	665
C. felis esterase,	AAV40727	19	401	7.7	14	c	666

Left column:

c	No.			Pos	Len	Accession	Description
c	667	7.7	14	401	19	AAV40728	C. felis esterase,
	668	7.7	14	401	22	AAD21163	Ctenocephalides fe
	669	7.7	14	405	23	ABV25055	Human prostate exp
	670	7.7	14	405	23	ABV37801	Human prostate exp
c	671	7.7	14	407	23	ABV37057	Human prostate exp
	672	7.7	14	408	23	ABV38242	Human prostate exp
	673	7.7	14	411	22	ABA44789	Human breast cell
	674	7.7	14	411	22	ABA55244	Human foetal liver
	675	7.7	14	411	22	ABA24991	Probe #3457 for ge
	676	7.7	14	411	22	AAK03502	Human brain expres
	677	7.7	14	411	22	AAI11542	Probe #3475 for ge
	678	7.7	14	411	22	AAI34904	Probe #3590 used t
	679	7.7	14	411	22	AAI03431	Probe #3322 used t
	680	7.7	14	411	24	ABS03488	Human genome-deriv
	681	7.7	14	412	22	AAK96620	Human neuregulin g
	682	7.7	14	412	22	AAK98113	Human neuregulin g
	683	7.7	14	413	22	ABA45881	Human breast cell
	684	7.7	14	413	22	ABA56403	Human foetal liver
	685	7.7	14	413	22	ABA26033	Probe #4499 for ge
	686	7.7	14	413	22	AAK04573	Human brain expres
	687	7.7	14	413	22	AAK30076	Human foetal liver
	688	7.7	14	413	22	AAI14670	Human bone marrow
	689	7.7	14	413	22	AAI36039	Probe #4603 for ge
c	690	7.7	14	413	22	AAI04481	Probe #4725 used t
c	691	7.7	14	413	24	ABS04652	Probe #4472 used t
c	692	7.7	14	415	22	ABA45430	Human genome-deriv
c	693	7.7	14	415	22	ABA55928	Human breast cell
c	694	7.7	14	415	22	ABA25595	Probe #4061 for ge
	695	7.7	14	415	22	AAK04142	Human brain expres
c	696	7.7	14	415	22	AAK29623	Human bone marrow
c	697	7.7	14	415	22	AAI14202	Probe #4135 for ge
c	698	7.7	14	415	22	AAI35583	Probe #4269 used t
c	699	7.7	14	415	22	AAI04039	Probe #4030 used t
	700	7.7	14	415	24	ABS04179	Human genome-deriv
	701	7.7	14	416	23	AAI93350	Human polynucleoti
c	702	7.7	14	424	23	ABV49874	Human prostate exp
c	703	7.7	14	426	18	AAT83769	DNA encoding a Sta
c	704	7.7	14	429	23	ABV17182	Human prostate exp
c	705	7.7	14	433	22	AAS58447	cDNA #1123 encodin
c	706	7.7	14	434	24	ABK54663	Human colon cancer
c	707	7.7	14	435	22	AAI79993	Human polynucleoti
c	708	7.7	14	444	22	AAF92326	Bovine mammary tis
c	709	7.7	14	444	24	ABL36587	Human colon tumour
c	710	7.7	14	448	23	ABA13937	Human nervous syst
c	711	7.7	14	448	22	ABN60966	Human cancer relat
c	712	7.7	14	448	24	ABL64670	Stomach cancer rel
c	713	7.7	14	450	22	ABA44797	Human breast cell
c	714	7.7	14	450	22	ABA55252	Human foetal liver
c	715	7.7	14	450	22	ABA59258	Human foetal liver
c	716	7.7	14	450	22	ABA24997	Probe #3463 for ge
c	717	7.7	14	450	22	ABA27997	Probe #6463 used t
c	718	7.7	14	450	24	AAI86704	Human polynucleoti
c	719	7.7	14	450	22	ABL36587	Human brain expres
c	720	7.7	14	450	22	AAK03508	Human brain expres
c	721	7.7	14	450	22	AAK07471	Human bone marrow
c	722	7.7	14	450	22	AAK33255	Human bone marrow
c	723	7.7	14	450	22	AAI13549	Probe #3482 for ge
c	724	7.7	14	450	22	AAI34911	Probe #3597 used t
c	725	7.7	14	450	22	AAI39049	Probe #7735 used t
c	726	7.7	14	450	22	AAI03438	Probe #3429 used t
c	727	7.7	14	450	24	ABS03496	Human genome-deriv
c	728	7.7	14	450	22	ABS08087	Human genome-deriv
c	729	7.7	14	456	22	ABA43700	Human breast cell
c	730	7.7	14	456	22	ABA54156	Human foetal liver
c	731	7.7	14	456	22	ABA23901	Probe #2367 for ge
c	732	7.7	14	456	22	ABA23901	Human brain expres
c	733	7.7	14	456	22	AAK02427	Human bone marrow
c	734	7.7	14	456	22	AAK27870	Probe #2382 for ge
c	735	7.7	14	456	22	AAI12449	Probe #2490 used t
c	736	7.7	14	456	22	AAI33804	Human genome-deriv
c	737	7.7	14	456	22	AAI02360	Human genome-deriv
c	738	7.7	14	457	24	ABS02327	Ovary cancer relat
c	739	7.7	14	458	21	ABL67850 / AAA30942	Human Expressed se

Right column:

Description	Accession	Len	Pos			No.	c
Human reproductive	AAL05585	22	458	7.7	14	740	c
Human lung tumour	AAF68129	22	459	7.7	14	741	c
cDNA encoding clon	ABK38040	24	459	7.7	14	742	c
Human colon cancer	ABQ58465	24	463	7.7	14	743	
Human prostate spe	AAH26848	22	465	7.7	14	744	c
Human colon tumour	ABL36439	24	465	7.7	14	745	
Human prostate exp	ABV46977	23	467	7.7	14	746	
Human breast cance	AAL09976	22	470	7.7	14	747	c
Human cardiovascul	ABA36433	22	470	7.7	14	748	
Human foetal liver	ABA74729	22	471	7.7	14	749	c
Probe #17906 for g	ABA39440	22	471	7.7	14	750	c
Probe #5189 for ge	AAI15256	22	471	7.7	14	751	c
Probe #16432 for g	AAI26499	22	471	7.7	14	752	c
Probe #23935 used	AAI55249	22	471	7.7	14	753	c
Probe #4981 used t	AAI04990	22	471	7.7	14	754	c
Nucleotide sequenc	AAI09911	22	473	7.7	14	755	c
Human genome-deriv	ABS22874	24	474	7.7	14	756	c
Human breast cance	AAL09911	22	474	7.7	14	757	c
Human brain expres	AAK04984	22	476	7.7	14	758	
Drosophila melanog	ABA28829	22	478	7.7	14	759	
Human foetal liver	ABA58280	22	478	7.7	14	760	
Human brain expres	AAK06374	22	478	7.7	14	761	c
Human bone marrow	AAK32042	22	478	7.7	14	762	c
Human genome-deriv	ABS06809	24	478	7.7	14	763	c
Myrtaceae microsat	AAA35313	21	480	7.7	14	764	c
DNA encoding novel	AAS27789	22	480	7.7	14	765	c
cDNA encoding nove	AAS27324	22	483	7.7	14	766	
Human polynucleoti	AAI63947	23	483	7.7	14	767	
Human prostate exp	ABV31850	23	483	7.7	14	768	c
Lung small cell ca	AAS61706	24	483	7.7	14	769	
Lung small cell ca	AAS61853	24	484	7.7	14	770	
Zea mays DNA fragm	AAC48122	21	484	7.7	14	771	
Human genome-deriv	ABV11119	23	487	7.7	14	772	
Human prostate exp	ABV47281	23	494	7.7	14	773	
Rat U3 gene trap d	AAX57489	20	498	7.7	14	774	
C. glutamicum SRT	AAF71049	22	501	7.7	14	775	
Murine transcripti	AAK54261	22	503	7.7	14	776	c
LEMMI 10 cDNA. Ly	AAQ32658	13	507	7.7	14	777	c
Human reproductive	AAQ12944	22	507	7.7	14	778	c
Human testicular a	AAL05096	22	507	7.7	14	779	
Human immune/haema	ABL97989	22	508	7.7	14	780	
Human cancer relat	AAK62779	22	510	7.7	14	781	
Human prostate exp	ABN63630	23	516	7.7	14	782	
DNA encoding novel	ABV47626	22	517	7.7	14	783	c
Rat beta amyloid f	AAK61428	23	520	7.7	14	784	
Human beta amyloid	ABK44559	16	520	7.7	14	785	
Human colon cancer	AAQ82909	24	520	7.7	14	786	
Human nervous syst	AAQ82910	22	522	7.7	14	787	
Human prostate exp	ABQ58332	23	522	7.7	14	788	
Human prostate exp	ABA11395	23	522	7.7	14	789	
Porcine muscular s	ABV39145	24	524	7.7	14	790	c
Human secreted pro	ABV45483	23	525	7.7	14	791	
Human immune/haema	AAS61953	22	527	7.7	14	792	
DNA encoding novel	AAC79029	22	527	7.7	14	793	
cDNA encoding colo	AAK61428	22	529	7.7	14	794	
Human foetal liver	ABK44559	22	529	7.7	14	795	
Probe #16322 for g	ABA71775	22	529	7.7	14	796	c
Human brain expres	ABA37856	22	529	7.7	14	797	
Human bone marrow	AAK20144	22	529	7.7	14	798	
Human bone marrow	AAK37214	22	530	7.7	14	799	
Probe #15498 for g	AAK46204	22	532	7.7	14	800	c
Probe #20796 used	AAI25565	22	533	7.7	14	801	c
Rat sequence diffe	AAI52110	24	534	7.7	14	802	c
Drosophila melanog	ABK62168	22	536	7.7	14	803	
Human colon cancer	AAH28827	24	537	7.7	14	804	c
Human cancer relat	ABQ58269	24	529	7.7	14	805	
Human cancer relat	ABN65694	22	529	7.7	14	806	
Murine 7-transmemb	AAC74496	22	534	7.7	14	807	
Human ORFX ORF2C51	AAF71048	21	536	7.7	14	808	c
C. glutamicum SRT	AAF71048	22	537	7.7	14	809	
Human prostate exp	ABV37112	23	542	7.7	14	810	c
Human lung tumour	AAF68173	22	547	7.7	14	811	
Human prostate exp	ABV01527	23	547	7.7	14	812	c

us-09-550-163-1_copy_80_260.oligo.rng

Seq	7.7	14	Num	Accession	Len	Description
813	7.7	14	547	ABK38084	24	cDNA encoding clon
814	7.7	14	549	AAC95226	21	Cat flea head and
815	7.7	14	552	ABA60637	22	Human foetal liver
816	7.7	14	552	ABA28744	22	Probe #7210 for ge
817	7.7	14	552	AAK08919	22	Human brain expres
818	7.7	14	552	AAK34809	24	Human bone marrow
819	7.7	14	552	AAI40528	22	Probe #9214 used t
820	7.7	14	552	ABQ09358	24	Human genome-deriv
821	7.7	14	557	ABQ44438	24	Oligonucleotide fo
822	7.7	14	557	ABQ44439	24	Oligonucleotide fo
c 823	7.7	14	560	ABV59436	23	Human prostate exp
824	7.7	14	561	AAH77920	22	Nucleotide sequenc
825	7.7	14	570	ABA63981	22	Human foetal liver
826	7.7	14	570	ABA31152	22	Human brain expres
827	7.7	14	570	AAK12478	22	Probe #9618 for ge
828	7.7	14	570	AAK38188	22	Human bone marrow
829	7.7	14	570	AAI88981	22	Probe #12805 used
830	7.7	14	570	AAI44119	24	Human genome-deriv
831	7.7	14	572	ABS12222	24	Human breast cance
c 832	7.7	14	572	AAL18495	22	Human prostate exp
833	7.7	14	572	ABV51541	23	Human bone marrow
c 834	7.7	14	573	AAK37952	22	cDNA encoding huma
835	7.7	14	578	AAA78135	21	Colon tumour relat
c 836	7.7	14	578	AAI28873	22	Human colon cancer
837	7.7	14	578	ABQ59207	22	Human breast cance
838	7.7	14	579	AAL09823	22	Human breast cance
839	7.7	14	580	AAS61567	24	Lung small cell ca
840	7.7	14	583	ABA62779	22	Human foetal liver
841	7.7	14	583	ABA30072	22	Probe #8538 for ge
842	7.7	14	583	AAK11177	22	Human brain expres
843	7.7	14	583	AAK36984	24	Human bone marrow
c 844	7.7	14	583	AAI17826	22	Probe #7759 for ge
845	7.7	14	583	AAI42796	22	Probe #11482 used
846	7.7	14	585	ABS10991	24	Human genome-deriv
847	7.7	14	592	AAX20391	20	Borrelia burgdorfe
c 848	7.7	14	592	AAX90814	22	Rat progression re
849	7.7	14	593	AAH11477	22	Human olfactory re
850	7.7	14	593	AAH32499	22	Human olfactory re
851	7.7	14	596	ABA63031	22	Human foetal liver
852	7.7	14	596	ABA30291	22	Probe #8757 for ge
853	7.7	14	596	AAI18064	22	Probe #7997 for ge
854	7.7	14	596	AAI41063	22	Probe #11749 used
855	7.7	14	597	ABS11222	24	Human genome-deriv
856	7.7	14	599	ABS15669	22	Human prostate exp
c 857	7.7	14	599	AAV43880	19	Mycobacterial meth
858	7.7	14	599	ABA62429	22	Human foetal liver
859	7.7	14	599	ABA29763	22	Probe #8229 for ge
860	7.7	14	599	AAS41978	24	Genomic sequence #
861	7.7	14	599	AAK10769	22	Human brain expres
c 862	7.7	14	600	ABA09192	22	Human MiRP1 homolo
c 863	7.7	14	601	ABS62645	22	Human bone marrow
864	7.7	14	604	AAI17498	22	Probe #7431 for ge
865	7.7	14	604	AAI42408	22	Probe #11094 used
866	7.7	14	605	AAK75620	24	Human genome-deriv
867	7.7	14	605	ABS10642	24	Human cancer relat
c 868	7.7	14	609	ABV32602	23	Human prostate exp
869	7.7	14	609	ABV41527	23	Human prostate exp
870	7.7	14	610	ABN68124	24	Streptococcus poly
871	7.7	14	610	ABK62626	24	Rat sequence diffe
c 872	7.7	14	611	ABQ55525	24	Human ovarian anti
873	7.7	14	620	AAS41134	22	cDNA encoding nove
874	7.7	14	620	AAS00311	22	S. pneumoniae expr
c 875	7.7	14	626	ABV46531	23	Human prostate exp
876	7.7	14	626	AAQ66151	24	Arabidopsis thalia
877	7.7	14	627	ABQ59915	22	Human colon cancer
878	7.7	14	627	AAX20950	20	Polynucleotide seq
c 879	7.7	14	634	AAH77918	22	Nucleotide sequenc
880	7.7	14	636	ABV36646	23	Human prostate exp
881	7.7	14	637	ABV37774	23	Human prostate exp

Seq	7.7	14	Num	Len	Accession	Description
c 886	7.7	14	637	23	ABV40144	Human prostate exp
c 887	7.7	14	637	23	ABV41331	Human prostate exp
c 888	7.7	14	638	23	AAS41933	Genomic sequence #
889	7.7	14	639	22	AAH52867	S. epidermidis ope
890	7.7	14	640	22	AAK56988	Human immune/haema
891	7.7	14	645	24	ABQ86213	Lactobacillus rham
892	7.7	14	649	22	ABQ59974	Human colon cancer
893	7.7	14	649	24	AAK64518	Human immune/haema
894	7.7	14	650	24	ABL04535	Drosophila melanog
895	7.7	14	651	23	ABQ58662	Human colon cancer
896	7.7	14	653	24	ABL89803	Human colon cancer
897	7.7	14	655	22	AAC93368	Human secreted pro
898	7.7	14	655	21	AAK51661	Human polynucleoti
c 899	7.7	14	657	22	ABV16733	Human prostate exp
900	7.7	14	661	24	AAI93354	Human polynucleoti
901	7.7	14	661	22	AAF12297	Aspergillus oryzae
c 902	7.7	14	669	23	AAF44921	Human breast cance
c 903	7.7	14	674	23	ABV22418	Human prostate exp
c 904	7.7	14	675	21	ABV28231	Human prostate exp
c 905	7.7	14	677	21	AAS62171	Porcine muscular s
c 906	7.7	14	687	21	AAC45257	Arabidopsis thalia
c 907	7.7	14	691	24	ABA90333	Human polynucleoti
c 908	7.7	14	695	24	ABQ70054	Human colon cancer
c 909	7.7	14	700	22	AAC69587	Listeria monocytog
910	7.7	14	700	24	AAD08206	Human secreted pro
911	7.7	14	703	24	ABK53572	Human eosinophil-m
912	7.7	14	704	20	AAI96882	Human neuroblastom
913	7.7	14	705	22	AAZ15444	Human gene express
c 914	7.7	14	708	14	AAQ37896	KpnI-SalI beta-cas
915	7.7	14	708	14	AAQ38098	Sequence encoding
c 916	7.7	14	711	19	AAV25591	Human beta-casein
c 917	7.7	14	713	20	AAX99611	Nucleic acid seque
c 918	7.7	14	715	24	ABQ57495	Human colon cancer
c 919	7.7	14	716	24	AAS62101	Porcine muscular s
c 920	7.7	14	720	20	AAZ17556	Human gene express
c 921	7.7	14	721	24	AAS62104	Porcine muscular s
c 922	7.7	14	724	16	AAT15105	Down-regulated sen
c 923	7.7	14	726	22	AAI96906	Human neuroblastom
c 924	7.7	14	727	22	AAI97863	Human neuroblastom
c 925	7.7	14	728	22	AAI97821	Human neuroblastom
c 926	7.7	14	729	22	AAK63925	Human immune/haema
c 927	7.7	14	732	22	ABK55412	Human colon cancer
c 928	7.7	14	732	21	AAC64071	Human potassium ch
c 929	7.7	14	732	21	AAC64083	Human potassium ch
c 930	7.7	14	732	21	AAC64084	Human potassium ch
c 931	7.7	14	732	21	AAC64085	Human potassium ch
c 932	7.7	14	732	21	AAC64086	Human potassium ch
c 933	7.7	14	735	24	ABK86573	cDNA encoding huma
c 934	7.7	14	736	24	AAD35169	Human KCNE2 mutant
c 935	7.7	14	737	24	AAD35170	Human KCNE2 wild t
c 936	7.7	14	738	24	AAD35171	Human KCNE2 mutant
c 937	7.7	14	739	24	AAD35172	Human KCNE2 mutant
c 938	7.7	14	740	24	AAD35173	Human KCNE2 mutant
c 939	7.7	14	741	22	AAH04230	Human cDNA clone (
c 940	7.7	14	747	22	AAF82503	Human breast tumou
c 941	7.7	14	747	24	AAZ17474	Human gene express
c 942	7.7	14	749	24	AAS61964	Porcine muscular s
c 943	7.7	14	750	20	AAC45792	Arabidopsis thalia
c 944	7.7	14	751	21	AAI97344	Human neuroblastom
945	7.7	14	754	22	AAS73198	DNA encoding novel
946	7.7	14	755	23	ABO65437	Arabidopsis thalia
947	7.7	14	759	24	ABN70596	Streptococcus poly
948	7.7	14	759	22	AAH00630	Enterococcus faeci
949	7.7	14	749	22	AAL05900	Human reproductive
950	7.7	14	751	24	AAL05902	Human reproductive
951	7.7	14	754	22	ABL98464	Human testicular a
952	7.7	14	754	23	ABL98466	Human colon cancer
953	7.7	14	754	22	AAD18837	Human testicular a
954	7.7	14	755	22	AAH67152	Human ovarian tumo
955	7.7	14	759	22	AAZ15749	C glutamicum codin
957	7.7	14	761	22	AAX98719	Human gene express
958	7.7	14	761	20		Human validated ca

959		14	7.7	765	ABN69768	Streptococcus poly
c 960		14	7.7	768	AAI97633	Human neuroblastom
961		14	7.7	769	AAL24801	Human breast cance
c 962		14	7.7	771	AAS62012	Porcine muscular s
963		14	7.7	778	AAC35265	Arabidopsis thalia
c 964		14	7.7	779	AAZ97278	Human prostate can
c 965		14	7.7	781	AAS62095	Porcine muscular s
c 966		14	7.7	783	AAH07079	Human cDNA clone (
c 967		14	7.7	784	AAS61903	Porcine muscular s
968		14	7.7	785	AAX57419	Rat U3 gene trap d
c 969		14	7.7	787	AAI97250	Human neuroblastom
c 970		14	7.7	789	AAZ17435	Human gene express
971		14	7.7	789	AAS61905	Porcine muscular s
c 972		14	7.7	803	AAS61925	Porcine muscular s
973		14	7.7	804	AAS55851	Streptococcus pneu
c 974		14	7.7	810	AAZ17059	Human gene express
975		14	7.7	811	AAZ97258	Human prostate can
c 976		14	7.7	814	AAL18648	Human breast cance
977		14	7.7	818	AAH07537	Human cDNA clone (
c 978		14	7.7	819	AAS62143	Porcine muscular s
979		14	7.7	831	AAZ97392	Human prostate can
c 980		14	7.7	831	AAS55383	Staphylococcus aur
c 981		14	7.7	836	AAL01738	Human reproductive
c 982		14	7.7	836	ABL97031	Human testicular a
c 983		14	7.7	837	AAZ97338	Human prostate can
984		14	7.7	838	ABQ68658	Listeria monocytog
985		14	7.7	838	ABQ70294	Listeria monocytog
986		14	7.7	840	AAQ79136	Bovine male enhanc
987		14	7.7	842	AAH07594	Human cDNA clone (
988		14	7.7	843	AAX98680	Human validated ca
c 989		14	7.7	856	AAZ77501	Human ovarian tumo
990		14	7.7	858	ABN71103	Streptococcus poly
c 991		14	7.7	863	AAH08135	Human cDNA clone (
c 992		14	7.7	867	AAF71319	Corynebacterium gl
993		14	7.7	871	AAQ79137	Bovine male enhanc
c 994		14	7.7	883	AAH99078	Human EST-derived
995		14	7.7	885	AAC36999	Arabidopsis thalia
c 996		14	7.7	891	ABK74081	Bacillus lichenifo
c 997		14	7.7	894	AAS64119	DNA encoding novel
c 998		14	7.7	897	ABN68564	Streptococcus poly
999		14	7.7	899	ABQ34104	Oligonucleotide fo
c1000		14	7.7	899	ABQ34105	Oligonucleotide fo

ALIGNMENTS

RESULT 1
AAI24432
ID   AAI24432 standard; DNA; 372 BP.
XX
AC   AAI24432;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #14365 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.

PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID NO 14365; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%;   Score 181;   DB 22;   Length 372;
Best Local Similarity   100.0%;   Pred. No. 5.4e-82;
Matches 181;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ACTTTATCCAATTTCACACAGACCGTCTTCGAACGACGATTTTTATTACTTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 ACTTTATCCAATTTCACACAGACCGTCTTCGAACGACGATTTTTATTACTTAT 66

Qy     61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 126

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 186

Qy    181 A 181
          |
Db    187 A 187
```

RESULT 2
AAI09965
ID   AAI09965 standard; DNA; 372 BP.
XX
AC   AAI09965;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #9956 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.

PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 9956; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%; Score 181; DB 22; Length 372;
Best Local Similarity   100.0%; Pred. No. 5.4e-82;
Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 ACTTTATCCAATTTCACACAGACGCTCTTCCGAAGGATTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ACTTTATCCAATTTCACACAGACGCTCTTCCGAAGGATTTTATTACTTAT 66

Qy         61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
              ||||||||||||||||||||||||||||
Db         67 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 126

Qy        121 GCTGAGAACTTCTACTACTGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
              ||||||||||||||||||||||||||||
Db        127 GCTGAGAACTTCTACTACTGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 186

Qy        181 A 181
              |
Db        187 A 187
```

RESULT 3
AAS00245
ID   AAS00245 standard; DNA; 372 BP.
XX
AC   AAS00245;
XX
DT   10-MAY-2001 (first entry)
XX
DE   Human potassium channel regulatory protein, Mink2, DNA sequence.
XX
KW   Human; Mink2; potassium channel; cardiac arrhythmia; hypertension; ds;
KW   angina; asthma; diabetes; renal insufficiency; urinary incontinence;
KW   irritable colon; epilepsy; cerebrovascular ischaemia; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..372
FT   /*tag= a
FT   /product= "MINK2 potassium channel protein"
XX

PN   WO200114403-A1.
XX
PD   01-MAR-2001.
XX
PF   18-AUG-2000; 2000WO-US22799.
XX
PR   20-AUG-1999; 99US-0379201.
XX
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI   Ficker E, Wible B, Brown AM;
XX
DR   WPI; 2001-218424/22.
DR   P-PSDB; AAU00215.
XX
PT   Novel potassium channel gene termed MinK2 encoding potassium channel
PT   regulatory protein, useful for screening compounds that are useful for
PT   treating diseases caused by aberrant potassium activity -
XX
PS   Claim 1; Fig 9; 39pp; English.
XX
CC   The sequence represents the coding sequence of human potassium channel
CC   regulatory protein, Mink2. Mink2 sequence is useful for producing a
CC   potassium channel regulatory protein useful for in vitro or in vivo
CC   screening of agonistic or antagonistic compounds that are useful for
CC   treating diseases caused by aberrant potassium activity, such as human
CC   cardiac arrhythmias, hypertension, angina, asthma, diabetes, renal
CC   insufficiency, urinary incontinence, irritable colon, epilepsy,
CC   cerebrovascular ischaemia, and autoimmune disease.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%; Score 181; DB 22; Length 372;
Best Local Similarity   100.0%; Pred. No. 5.4e-82;
Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 ACTTTATCCAATTTCACACAGACGCTCTTCCGAAGGATTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ACTTTATCCAATTTCACACAGACGCTCTTCCGAAGGATTTTATTACTTAT 66

Qy         61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
              ||||||||||||||||||||||||||||
Db         67 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 126

Qy        121 GCTGAGAACTTCTACTACTGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
              ||||||||||||||||||||||||||||
Db        127 GCTGAGAACTTCTACTACTGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 186

Qy        181 A 181
              |
Db        187 A 187
```

RESULT 4
ABA44797
ID   ABA44797 standard; DNA; 450 BP.
XX
AC   ABA44797;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3492.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PP   30-JAN-2001; 2001WO-US00662.

```
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes -
XX
PS   Claim 1; SEQ ID NO 3492; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match             100.0%;  Score 181;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 5.4e-82;
Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 221

Qy     61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 281

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 341

Qy    181 A 181
          |
Db    342 A 342


RESULT 5
ABA55252
ID   ABA55252 standard; DNA; 450 BP.
XX
AC   ABA55252;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3557.
XX
```

```
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 3557; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match             100.0%;  Score 181;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 5.4e-82;
Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 221

Qy     61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 281

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 341

Qy    181 A 181
          |
Db    342 A 342


RESULT 6
ABA24997
ID   ABA24997 standard; DNA; 450 BP.
XX
AC   ABA24997;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #3463 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
```

Left column:

```
KW  cardiovascular disease; hypertension; cardiac arrhythmia;
KW  congenital heart disease; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157274-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00666.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-488899/53.
XX
PT  Single exon nucleic acid probes for analyzing gene expression in human
PT  hearts -
XX
PS  Claim 1; SEQ ID No 3463; 530pp; English.
XX
CC  The present invention relates to single exon nucleic acid probes for
CC  measuring human gene expression in a sample derived from human heart. The
CC  present sequence is one such probe. The probes may be used for
CC  predicting, measuring and displaying gene expression in samples derived
CC  from the human heart via microarrays. By measuring gene expression, the
CC  probes are useful for predicting, diagnosing, grading, staging,
CC  monitoring and prognosing diseases of the human heart and vascular system
CC  e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC  congenital heart disease.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;
```

Query Match 100.0%; Score 181; DB 22; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.4e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  162 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGGACAATTGGCGCCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 281

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGATGATTGGAATGTTCTCTTTC 341

Qy  181 A 181
        |
Db  342 A 342
```

```
RESULT 7
AAK03508
ID  AAK03508 standard; DNA; 450 BP.
XX
AC  AAK03508;
XX
DT  05-NOV-2001 (first entry)
```

Right column:

```
DE  Human brain expressed single exon probe SEQ ID NO: 3499.
XX
KW  Human; brain expressed exon; gene expression analysis; probe;
KW  microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW  epilepsy; cancer; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157275-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00667.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-483446/52.
XX
PT  Single exon nucleic acid probes for analyzing gene expression in human
PT  brains -
XX
PS  Example 4; SEQ ID NO: 3499; 650pp + Sequence Listing; English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in the human
CC  brain. They can be used to measure gene expression in brain cell samples,
CC  which may enable the diagnosis and improved treatment of nervous system
CC  diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC  epilepsy and cancers. The present sequence is one of the probes of the
CC  invention.
XX
SQ  Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;
```

Query Match 100.0%; Score 181; DB 22; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.4e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  162 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGGACAATTGGCGCCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 281

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGATGATTGGAATGTTCTCTTTC 341

Qy  181 A 181
        |
Db  342 A 342
```

```
RESULT 8
AAK28962
ID  AAK28962 standard; DNA; 450 BP.
XX
AC  AAK28962;
XX
DT  06-NOV-2001 (first entry)
XX
```

DE   Human bone marrow expressed single exon probe SEQ ID NO: 3519.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.
PR   21-SEP-2000; 2000US-0236359.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 3519; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match   100.0%;   Score 181;   DB 22;   Length 450;
Best Local Similarity 100.0%;   Pred. No. 5.4e-82;
Matches 181;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAGTTGAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAGTTGAT 281

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 341

Qy   181 A 181
         |
Db   342 A 342

RESULT 9
AAI13549
ID   AAI13549 standard; DNA; 450 BP.
XX
AC   AAI13549;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #3482 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;

KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID NO 3482; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification. but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match   100.0%;   Score 181;   DB 22;   Length 450;
Best Local Similarity 100.0%;   Pred. No. 5.4e-82;
Matches 181;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAAGAGGCCCTCCAAGCCAAGTTGAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAAGAGGCCCTCCAAGCCAAGTTGAT 281

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 341

Qy   181 A 181
         |
Db   342 A 342

RESULT 10
AAI34911
ID   AAI34911 standard; DNA; 450 BP.
XX
AC   AAI34911;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #3597 used to measure gene expression in human placenta sample.
XX

```
KW  Probe; microarray; human; placenta; antenatal diagnosis;
KW  genetic disorder; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157272-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00663.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-488897/53.
XX
PT  Human genome-derived single exon nucleic acid probes useful for
PT  analyzing gene expression in human placenta -
XX
PS  Claim 25; SEQ ID No 3597; 654pp; English.
XX
CC  The present invention relates to single exon nucleic acid probes (SENP).
CC  The present sequence is one such probe. The probes are useful for
CC  producing a microarray for predicting, measuring and displaying gene
CC  expression in samples derived from human placenta. The probes are useful
CC  for antenatal diagnosis of human genetic disorders.
XX
SQ  Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match           100.0%;  Score 181;  DB 22;  Length 450;
Best Local Similarity 100.0%;  Pred. No. 5.4e-82;
Matches 181;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT  221

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  281

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC  341

Qy  181 A  181
        |
Db  342 A  342

RESULT 11
AAI03438
ID  AAI03438 standard; DNA; 450 BP.
XX
AC  AAI03438;
XX
DT  09-OCT-2001  (first entry)
XX
DE  Probe #3429 used to measure gene expression in human breast sample.
XX
KW  Probe; human; breast disease; breast cancer; development disorder; ss;
KW  inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS  Homo sapiens.
```

```
XX
PN  WO200157270-A2.
XX
PD  09-AUG-2001.
XX
PF  29-JAN-2001; 2001WO-US00661.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-476286/51.
XX
PT  Novel single exon nucleic acid probe used to measuring gene expression
PT  in a human breast -
XX
PS  Claim 25; SEQ ID No 3429; 322pp; English.
XX
CC  The present invention relates to novel single exon nucleic acid probes.
CC  The present sequence is one such probe. The probes are useful for
CC  measuring human gene expression in a human breast sample, where the probe
CC  hybridies at high stringency to a nucleic acid expressed in the human
CC  breast. The probes are useful for predicting, diagnosing, grading,
CC  staging, monitoring and prognosing diseases of the human breast,
CC  particularly those diseases with polygenic aetiology. The diseases
CC  include: breast cancer, disorders of development, inflammatory diseases
CC  of the breast, fibrocystic changes, proliferative breast disease and
CC  non-carcinoma tumours.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match           100.0%;  Score 181;  DB 22;  Length 450;
Best Local Similarity 100.0%;  Pred. No. 5.4e-82;
Matches 181;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT  221

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  222 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  281

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC  341

Qy  181 A  181
        |
Db  342 A  342

RESULT 12
ABS03496
ID  ABS03496 standard; DNA; 450 BP.
XX
AC  ABS03496;
XX
DT  19-AUG-2002  (first entry)
XX
DE  Human genome-derived single exon probe from lung SEQ ID No 3487.
XX
```

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW primary ciliary dyskinesis; fibrocyutic pulmonary dysplasia;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 1; SEQ ID No 3487; 634pp; English.
XX
CC The invention relates to a spatially-addreseable set of single exon
CC nucleic acid probes for measuring gene uxpression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucloic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array, identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from gunomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridiuation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the spicification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak uyndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC pulmonary alveolar proteinosis, Karagenur syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe of the invention.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

    Query Match   100.0%; Score 181; DB 24; Length 450;
    Best Local Similarity 100.0%; Pred. No. 5.4e-82;
    Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 281

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 341

Qy   181 A 181
            |
Db   342 A 342

RESULT 13
AAI15256
ID AAI15256 standard; DNA; 471 BP.
XX
AC AAI15256;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #5189 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO20157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID No 5189; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene

```
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 471 BP; 148 A; 102 C; 86 G; 135 T; 0 other;

Query Match           100.0%;  Score 181;  DB 22;  Length 471;
Best Local Similarity 100.0%;  Pred. No. 5.4e-82;
Matches  181;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      245 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 304

Qy       61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      305 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 364

Qy      121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      365 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 424

Qy      181 A 181
            |
Db      425 A 425

RESULT 14
AAI04990
ID   AAI04990 standard; DNA; 471 BP.
XX
AC   AAI04990;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #4981 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 4981; 322pp; English.
XX
CC   The present invention relates to novel single nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
```

```
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 471 BP; 148 A; 102 C; 86 G; 135 T; 0 other;

Query Match           100.0%;  Score 181;  DB 22;  Length 471;
Best Local Similarity 100.0%;  Pred. No. 5.4e-82;
Matches  181;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      245 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 304

Qy       61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      305 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 364

Qy       61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      305 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 364

Qy      121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      365 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 424

Qy      181 A 181
            |
Db      425 A 425

RESULT 15
AAF80269
ID   AAF80269 standard; DNA; 471 BP.
XX
AC   AAF80269;
XX
DT   29-JUN-2001 (first entry)
XX
DE   Nucleotide sequence of human potassium channel subunit IsK2.
XX
KW   Human; potassium channel; IsK2; gene therapy; gastric motility;
KW   gastric acid secretion; anti-arrhythmic agent; myocardial infarction; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             79..450
FT                   /*tag= a
FT                   /product= "potassium channel subunit IsK2"
XX
PN   WO200127246-A1.
XX
PD   19-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US28014.
XX
PR   12-OCT-1999;  99US-0158781.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Swanson RJ, Liu Y, Folander K;
XX
DR   WPI; 2001-273764/28.
DR   P-PSDB; AAB67800.
XX
PT   New DNA encoding the IsK2 potassium channel subunit, useful e.g. for
PT   detecting mutations and screening for therapeutic agents -
XX
PS   Claim 3; Fig 1A; 46pp; English.
XX
```

CC The present sequence encodes a human potassium channel subunit,
CC designated IsK2. The IsK2 polynucleotide, and derived probes, are
CC used diagnostically to detect mutations in the IsK2 gene, to determine
CC levels of mRNA expression and to isolate homologous sequences; for
CC recombinant expression of IsK2; in gene therapy to increase potassium
CC channel activity and to generate transgenic animals, as models and
CC for drug screening. Recombinant IsK2 is used for studying biochemical
CC activity of IsK2 and its role in disorders of gastric motility and
CC gastric acid secretion, and to raise specific antibodies. IsK2
CC modulators are potentially useful for treating diseases associated with
CC increased or reduced potassium channel activity, e.g. as
CC anti-arrhythmic agents for treating myocardial infarction and as
CC regulators of gastric acid secretion.
XX
SQ Sequence 471 BP; 143 A; 110 C; 103 G; 115 T; 0 other;

Query Match   100.0%; Score 181; DB 22; Length 471;
Best Local Similarity 100.0%; Pred. No. 5.4e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 144

Qy   61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 204

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 264

Qy   181 A 181
          |
Db   265 A 265

RESULT 16
ABA09192
ID ABA09192 standard; cDNA; 600 BP.
XX
AC ABA09192;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human MiRP1 homologue-encoding cDNA, SEQ ID NO:968.
XX
OS Homo sapiens.
XX
PN WO20015718B-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder; immune disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.

XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11948.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 826; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 600 BP; 187 A; 133 C; 130 G; 144 T; 6 other;

Query Match   100.0%; Score 181; DB 22; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.5e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   44 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 103

Qy   61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 163

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   164 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 223

Qy   181 A 181
          |
Db   224 A 224

RESULT 17
AAK52645
ID   AAK52645 standard; cDNA; 600 BP.
XX
AC   AAK52645;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 2174.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US04098.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   01-SEP-2000; 2000US-0654936.
PR   15-SEP-2000; 2000US-0663561.
PR   20-OCT-2000; 2000US-0693325.
PR   30-NOV-2000; 2000US-0728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,   Liu C,   Drmanac RT,   Asundi V,   Zhou P,   Xu C,   Cao Y,   Ma Y;
PI   Zhao QA,   Wang D,   Wang J,   Zhang J,   Ren F,   Chen R,   Wang ZW;
PI   Xue AJ,   Yang Y,   Wejhrman T,   Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM79512.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities,
PT   useful in diagnosis and gene therapy -
XX
PS   Claim 1; Page 4539-4540; 622lpp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC   (AAM80020) are omitted as the relevant pages from the sequence listing
CC   were missing at the time of publication.
XX
SQ   Sequence 600 BP; 187 A; 133 C; 130 G; 144 T; 6 other;

Query Match   100.0%; Score 181; DB 22; Length 600;
Best Local Similarity   100.0%; Pred. No. 5.5e-82;
Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   44 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 103

Qy   61 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAGAAGAGGCCCTCCAAGCCAAGTTGAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   104 ATGGACAATTGGCGCCAGAACACACAGCTGAGCAGAAGAGGCCCTCCAAGCCAAGTTGAT 163

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGTGATGATTGGAATGTTCTCTTC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   164 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGTGTGATGATTGGAATGTTCTCTTC 223

Qy   181 A 181
              |
Db   224 A 224

RESULT 18
AAK51661
ID   AAK51661 standard; cDNA; 655 BP.
XX
AC   AAK51661;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 206.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US04098.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   01-SEP-2000; 2000US-0654936.
PR   15-SEP-2000; 2000US-0663561.
PR   20-OCT-2000; 2000US-0693325.
PR   30-NOV-2000; 2000US-0728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,   Liu C,   Drmanac RT,   Asundi V,   Zhou P,   Xu C,   Cao Y,   Ma Y;
PI   Zhao QA,   Wang D,   Wang J,   Zhang J,   Ren F,   Chen R,   Wang ZW;
PI   Xue AJ,   Yang Y,   Wejhrman T,   Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM78528.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities,
PT   useful in diagnosis and gene therapy -
XX
PS   Claim 1; Page 1024; 622lpp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC   (AAM80020) are omitted as the relevant pages from the sequence listing
CC   were missing at the time of publication.
XX
SQ   Sequence 655 BP; 196 A; 154 C; 146 G; 153 T; 6 other;

Query Match   100.0%; Score 181; DB 22; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.5e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 158

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 218

Qy 121 GCTGAGGAACTTCTACTATGTCATCCTGTACCTCGTGGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 219 GCTGAGGAACTTCTACTATGTCATCCTGTACCTCGTGGTGATGATTGGAATGTTCTCTTTC 278

Qy 181 A 181
        |
Db 279 A 279

RESULT 19
AAC64071
ID AAC64071 standard; cDNA; 732 BP.
XX
AC AAC64071;
XX
DT 19-FEB-2001 (first entry)
XX
DE Human potassium channel protein KCNE2 (MiRP1) cDNA, SEQ ID NO:1.
XX
KW Human; KCNE2; MiRP1; potassium channel protein; KNCE1-related;
KW MinK-related; long QT syndrome; cardiac arrhythmia;
KW drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW HERG; ss.
XX
OS Homo sapiens.
XX
PN WO20006343434-A1.
XX
PD 26-OCT-2000.
XX
PF 14-APR-2000; 2000WO-US10004.
XX
PR 15-APR-1999; 99US-0129404.
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (UYYA) UNIV YALE.
XX
PI Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR WPI; 2000-672747/65.
DR P-PSDB; AAB29585.
XX
PT Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT diagnosing and treating ion channel disorders, especially long QT
PT syndrome -
PS Claim 1; Page 118-119; 132pp; English.
XX
CC The invention relates to novel ion channel proteins related to
CC KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC respectively). The cDNAs encoding these proteins are given in AAC64071-
CC AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC potassium channels (I-KR), mutations in which are associated with long
CC QT syndrome. The invention also relates to methods of diagnosing long QT
CC syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic

CC nonhuman animals comprising a heterologous ion channel protein gene
CC of the invention, a transgenic animal comprising human KNCE2 and HERG
CC DNA, and methods of and screening drugs for treating long QT syndrome
CC using KCNE2 proteins (including mutants), nucleic acids encoding them
CC and antibodies against KNCE2 proteins. The methods, antibodies, nucleic
CC acids, and proteins may be used for diagnosing or treating ion channel
CC disorders, especially long QT syndrome. Transgenic animals comprising
CC KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC The present sequence represents cDNA encoding human KCNE2 (MiRP1).
XX
SQ Sequence 732 BP; 221 A; 152 C; 157 G; 202 T; 0 other;

Query Match   100.0%; Score 181; DB 21; Length 732;
Best Local Similarity 100.0%; Pred. No. 5.5e-82;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy 121 GCTGAGGAACTTCTACTATGTCATCCTGTACCTCGTGGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 200 GCTGAGGAACTTCTACTATGTCATCCTGTACCTCGTGGTGATGATTGGAATGTTCTCTTTC 259

Qy 181 A 181
        |
Db 260 A 260

RESULT 20
ABK86573
ID ABK86573 standard; DNA; 732 BP.
XX
AC ABK86573;
XX
DT 24-SEP-2002 (first entry)
XX
DE cDNA encoding human ether-a-go-go related interacting protein MiRP1.
XX
KW Human; human ether-a-go-go related gene; HERG; KCR1; MiRP1;
KW long QT syndrome; LQT; single nucleotide polymorphism; cardiac arrythmia;
KW potassium channel; ss; gene.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   74..445
FT   /*tag= a
FT   /product= "MiRP1"
XX
PN WO200242735-A2.
XX
PD 30-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-US45644.
XX
PR 30-OCT-2000; 2000US-244340P.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Balser JR, George AL, Roden DM;
XX
DR WPI; 2002-527650/56.
DR P-PSDB; AAU99168.
XX
PT Identifying a potassium channel activity modulator for drug design,
PT comprises contacting a compound with a potassium channel and rat
PT cerebellar cDNA library (KCR1) polypeptide, and determining activity -
XX

```
XX
FH   Key             Location/Qualifiers
FT   CDS             74..445
FT                   /*tag= a
FT                   /*product= "Human MiRP1 mutant protein"
FT   variation       replace (420, C)
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN   WO200222875-A2.
XX
PD   21-MAR-2002.
XX
PF   11-SEP-2001; 2001WO-US28332.
XX
PR   11-SEP-2000; 2000US-231571P.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Goldstein SAN;
XX
DR   WPI; 2002-362360/39.
DR   P-PSDB; AAE22094.
XX
PT   Novel gene encoding Min-K related ion channel protein subunit and
PT   polymorphisms in this gene associated with antibiotic-induced long QT
PT   syndrome -
XX
PS   Claim 1; Page 41-42; 49pp; English.
XX
CC   The present invention relates to novel KCNE2 genes encoding Min-K related
CC   (MiRP) ion channel proteins and polymorphisms in these genes that are
CC   associated with ion channel disorders including antibiotic-induced long
CC   QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC   57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC   encoding the amino acid positions is useful for diagnosing the presence
CC   of a polymorphism that causes drug-induced LQTS. The diagnostic methods
CC   are useful in the development of new drug therapies which selectively
CC   target one or more KCNE2 polymorphisms that are associated with cardiac
CC   arrhythmias. The present sequence is human KCNE2 mutant DNA (C420T).
XX
SQ   Sequence 732 BP; 221 A; 151 C; 157 G; 203 T; 0 other;

Query Match               100.0%;  Score 181;  DB 24;  Length 732;
Best Local Similarity     100.0%;  Pred. No. 5.5e-82;
Matches  181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 199

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 199

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 259

Qy  181 A 181
        |
Db  260 A 260

RESULT 22
AAD35170
ID   AAD35170 standard; DNA; 732 BP.
XX
AC   AAD35170;
XX
DT   25-JUL-2002 (first entry)
XX
DE   Human KCNE2 wild type DNA.
```

```
PS   Claim 17; Page 162-163; 164pp; English.
XX
CC   The invention relates to identifying (M1) a compound that modulates
CC   biological activity of a potassium channel (PC), by contacting a
CC   compound with a structure comprising a PC polypeptide and a polypeptide
CC   cloned from a rat cerebellar cDNA library (KCR1), and determining the
CC   activity of the PC polypeptide in the presence and absence of the
CC   compound, where a difference in the activities indicates modulation of
CC   biological activity of PC. Also include are identifying (M2) a candidate
CC   compound that modulates the biological activity of a complex comprising a
CC   human ether-a-go-go-related gene (HERG) channel polypeptide and a KCR1
CC   polypeptide, identifying (M3) a candidate compound as a modulator of KCR1
CC   expression, modulating (M4) PC function in a subject, comprising
CC   administering to the subject a substance that provides expression of a
CC   KCR1-encoding nucleic acid molecule in a cell or tissue, where modulated
CC   PC function is desired, screening (M5) for susceptibility to a drug-
CC   induced cardiac arrhythmia in a subject, comprising obtaining a
CC   biological sample from the subject and detecting a polymorphism of a KCR1
CC   gene in the biological sample from the subject, where the presence of the
CC   polymorphism indicates the susceptibility of the subject to a
CC   drug-induced cardiac arrhythmia, an oligonucleotide pair, where a first
CC   oligonucleotide of the pair hybridises to a first portion of a KCR1 gene
CC   which includes a polymorphism of the KCR1 gene, and the second
CC   oligonucleotide of the pair hybridises to a second portion of the KCR1
CC   gene that is adjacent to the first portion and a set of antisense
CC   oligonucleotide primers, suitable for amplifying a portion of a KCR1 gene
CC   which includes a polymorphism of the KCR1 gene. (M1) is useful for
CC   identifying a compound that modulates biological activity of PC,
CC   especially HERG, for modulating PC function (i.e modulating HERG
CC   activity) in a mammal, by preparing a composition comprising the
CC   compound and administering the composition. The compound is useful for
CC   treating or preventing long QT syndrome (LQT) and is useful in drug
CC   designing. The present sequence encodes a HERG interacting
CC   protein MiRP1 (not defined).
XX
SQ   Sequence 732 BP; 221 A; 153 C; 157 G; 201 T; 0 other;

Query Match               100.0%;  Score 181;  DB 24;  Length 732;
Best Local Similarity     100.0%;  Pred. No. 5.5e-82;
Matches  181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGACGCCCTCCAAGCAAAGTTGAT 199

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 259

Qy  181 A 181
        |
Db  260 A 260

RESULT 21
AAD35169
ID   AAD35169 standard; DNA; 732 BP.
XX
AC   AAD35169;
XX
DT   25-JUL-2002 (first entry)
XX
DE   Human KCNE2 mutant DNA (C420T).
XX
KW   Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW   KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; mutant; gene; SNP;
KW   single nucleotide polymorphism; ds.
XX
OS   Homo sapiens.
```

```
XX
KW   Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW   KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             74..445
FT                   /tag= a
FT                   /product= "Human MiRP1 protein"
XX
PN   WO20022875-A2.
XX
PD   21-MAR-2002.
XX
PF   11-SEP-2001; 2001WO-US28332.
XX
PR   11-SEP-2000; 2000US-231571P.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Goldstein SAN;
XX
DR   WPI; 2002-362360/39.
DR   P-PSDB; AAE22095.
XX
PT   Novel gene encoding Min-K related ion channel protein subunit and
PT   polymorphisms in this gene associated with antibiotic-induced long QT
PT   syndrome -
XX
PS   Claim 9; Page 43; 49pp; English.
XX
CC   The present invention relates to novel KCNE2 genes encoding Min-K related
CC   (MiRP) 1 ion channel proteins and polymorphisms in these genes that are
CC   associated with ion channel disorders including antibiotic-induced long
CC   QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC   57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC   encoding the amino acid positions is useful for diagnosing the presence
CC   of a polymorphism that causes drug-induced LQTS. The diagnostic methods
CC   are useful in the development of new drug therapies which selectively
CC   target one or more KCNE2 polymorphisms that are associated with cardiac
CC   arrhythmias. The present sequence is human KCNE2 wild type DNA.
XX
SQ   Sequence 732 BP; 221 A; 152 C; 157 G; 202 T; 0 other;

Query Match             100.0%;  Score 181;  DB 24;  Length 732;
Best Local Similarity   100.0%;  Pred. No. 5.5e-82;
Matches  181;  Conservative   0;  Mismatches   0;  Indels    0;  Gaps    0;

Qy           1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTTATTACTTAT 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTTATTACTTAT 139

Qy          61 ATGGACAAATTGGCGCCAGAACAACACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         140 ATGGACAAATTGGCGCCAGAACAACACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy         121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTC 259

Qy         181 A 181
               |
Db         260 A 260
```

```
XX
DE   Human potassium channel protein KCNE2 (MiRP1) T8A mutant DNA.
XX
KW   Human; KCNE2; MiRP1; potassium channel protein; KNCE1-related;
KW   MinK-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW   fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW   HERG; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200063434-A1.
XX
PD   26-OCT-2000.
XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999;  99US-0129404.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (UYYA ) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672747/65.
DR   P-PSDB; AAB29596.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Claim 56; Page -; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC   respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KCNE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents DNA encoding a mutant human KCNE2
CC   (MiRP1) specifically claimed for use in diagnostic and drug screening
CC   methods of the invention.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the wild-type human KCNE2 cDNA sequence shown on page
CC   118-119.
XX
SQ   Sequence 732 BP; 220 A; 152 C; 158 G; 202 T; 0 other;

Query Match             91.2%;  Score 165;  DB 21;  Length 732;
Best Local Similarity   100.0%;  Pred. No. 7.1e-74;
Matches  165;  Conservative   0;  Mismatches   0;  Indels    0;  Gaps    0;

Qy          17 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTTATTACTTATATGGACAATTGGCGGCC 76
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTTATTACTTATATGGACAATTGGCGGCC 155

Qy          77 AGACACAACACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT 136
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         156 AGACACAACACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT 215
```

```
RESULT 23
AAC64086
ID   AAC64086 standard; DNA; 732 BP.
XX
AC   AAC64086;
XX
DT   19-FEB-2001 (first entry)
```

us-09-550-163-1_copy_80_260.oligo.rng

Qy   137 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTCA 181
          ||||||||||||||||||||||||||||||||||||||||||||
Db   216 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTCA 260

RESULT 24
AAD35173
ID   AAD35173 standard; DNA; 732 BP.
XX
AC   AAD35173;
XX
DT   25-JUL-2002 (first entry)
XX
DE   Human KCNE2 mutant DNA (A95G).
XX
KW   Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW   KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; mutant; gene; SNP;
KW   single nucleotide polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   74..445
FT   /tag= a
FT   variation   replace (95, A)
FT   /product= "Human MiRP1 mutant protein"
FT   /tag= b
FT   /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN   WO20022875-A2.
XX
PD   21-MAR-2002.
XX
PF   11-SEP-2001; 2001WO-US28332.
XX
PR   11-SEP-2000; 2000US-231571P.
XX
PA   (UYYA) UNIV YALE.
XX
PI   Goldstein SAN;
XX
DR   WPI; 2002-362360/39.
DR   P-PSDB; AAE22098.
XX
PT   Novel gene encoding Min-K related ion channel protein subunit and
PT   polymorphisms in this gene associated with antibiotic-induced long QT
PT   syndrome -
XX
PS   Claim 18; Page 47-48; 49pp; English.
XX
CC   The present invention relates to novel KCNE2 genes encoding Min-K related
CC   (MiRP) 1 ion channel proteins and polymorphisms in these genes that are
CC   associated with ion channel disorders including antibiotic-induced long
CC   QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC   57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC   encoding the amino acid positions is useful for diagnosing the presence
CC   of a polymorphism that causes drug-induced LQTS. The diagnostic methods
CC   are useful in the development of new drug therapies which selectively
CC   target one or more KCNE2 polymorphisms that are associated with cardiac
CC   arrhythmias. The present sequence is human KCNE2 mutant DNA (A95G).
XX
SQ   Sequence 732 BP; 220 A; 152 C; 158 G; 202 T; 0 other;

Query Match   91.2%; Score 165; DB 24; Length 732;
Best Local Similarity 100.0%; Pred. No. 7.1e-74;
Matches 165; Conservative 0; Mismatches 0; Indels   0; Gaps   0;

Qy   17 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTATATGACAATTGGCGCC 76
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   96 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTATATGACAATTGGCGCC 155

Qy   77 AGAACACAACAGTCGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT 136
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   156 AGAACACAACAGTCGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT 215

Qy   137 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTCA 181
         ||||||||||||||||||||||||||||||||||||||||||||
Db   216 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTCA 260

RESULT 25
AAC64085
ID   AAC64085 standard; DNA; 732 BP.
XX
AC   AAC64085;
XX
DT   19-FEB-2001 (first entry)
XX
DE   Human potassium channel protein KCNE2 (MiRP1) I57T mutant DNA.
XX
KW   Human; KCNE2; MiRP1; potassium channel protein; KCNE1-related;
KW   MinK-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW   fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW   HERG; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200063434-A1.
XX
PD   26-OCT-2000.
XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999;   99US-0129404.
XX
PA   (UTAH) UNIV UTAH RES FOUND.
PA   (UYYA) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672747/65.
DR   P-PSDB; AAB29595.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Claim 56; Page -; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC   respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KCNE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents DNA encoding a mutant human KCNE2
CC   (MiRP1) specifically claimed for use in diagnostic and drug screening
CC   methods of the invention.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the wild-type human KNCE2 cDNA sequence shown on page

```
CC  118-119.
XX
SQ  Sequence 732 BP; 221 A; 153 C; 157 G; 201 T; 0 other;
XX

Query Match     90.1%;  Score 163;  DB 21;  Length 732;
Best Local Similarity   100.0%;  Pred. No. 7.3e-73;
Matches 163;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 139

Qy  61 ATGGACAATTGGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGGACAATTGGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy 121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 163
        ||||||||||||||||||||||||||||||||||||||||||
Db 200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 242

RESULT 26
AAD35172
ID  AAD35172 standard; DNA; 732 BP.
XX
AC  AAD35172;
XX
DT  25-JUL-2002  (first entry)
XX
DE  Human KCNE2 mutant DNA (T243C).
XX
KW  Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW  KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; mutant; gene; SNP;
KW  single nucleotide polymorphism; ds.
XX
OS  Homo sapiens.
XX
FH  Key         Location/Qualifiers
FT  CDS         74..445
FT              /*tag= a
FT              /product= "Human MiRP1 mutant protein"
FT  variation   replace (243, T)
FT              /*tag= b
FT              /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN  WO200222875-A2.
XX
PD  21-MAR-2002.
XX
PF  11-SEP-2001; 2001WO-US28332.
XX
PR  11-SEP-2000; 2000US-231571P.
XX
PA  (UYYA ) UNIV YALE.
XX
PI  Goldstein SAN;
XX
DR  WPI; 2002-362360/39.
DR  P-PSDB; AAE22097.
XX
PT  Novel gene encoding Min-K related ion channel protein subunit and
PT  polymorphisms in this gene associated with antibiotic-induced long QT
PT  syndrome -
XX
PS  Claim 15; Page 46; 49pp; English.
XX
CC  The present invention relates to novel KCNE2 genes encoding Min-K related
CC  (MiRP) 1 ion channel proteins and polymorphisms in these genes that are
CC  associated with ion channel disorders including antibiotic-induced long
CC  QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC  57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC  encoding the amino acid positions is useful for diagnosing the presence
CC  of a polymorphism that causes drug-induced LQTS. The diagnostic methods
```

```
CC  are useful in the development of new drug therapies which selectively
CC  target one or more KCNE2 polymorphisms that are associated with cardiac
CC  arrhythmias. The present sequence is human KCNE2 mutant DNA (T243C).
XX
SQ  Sequence 732 BP; 221 A; 153 C; 157 G; 201 T; 0 other;

Query Match     90.1%;  Score 163;  DB 24;  Length 732;
Best Local Similarity   100.0%;  Pred. No. 7.3e-73;
Matches 163;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 139

Qy  61 ATGGACAATTGGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGGACAATTGGGCGCCAGAACACACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy 121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 163
        ||||||||||||||||||||||||||||||||||||||||||
Db 200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 242

RESULT 27
AAC64083
ID  AAC64083 standard; DNA; 732 BP.
XX
AC  AAC64083;
XX
DT  19-FEB-2001  (first entry)
XX
DE  Human potassium channel protein KCNE2 (MiRP1) Q9E mutant DNA.
XX
KW  Human; KCNE2; MiRP1; potassium channel protein; KNCE1-related;
KW  MinK-related; long QT syndrome; cardiac arrhythmia;
KW  drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW  fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW  HERG; mutant; ds.
XX
OS  Homo sapiens.
OS  Synthetic.
XX
PN  WO200063434-A1.
XX
PD  26-OCT-2000.
XX
PF  14-APR-2000; 2000WO-US10004.
XX
PR  15-APR-1999;  99US-0129404.
XX
PA  (UTAH ) UNIV UTAH RES FOUND.
PA  (UYYA ) UNIV YALE.
XX
PI  Abbott GW,  Sesti F,  Splawski I,  Keating MT,  Goldstein SAN;
XX
DR  WPI; 2000-672747/65.
DR  P-PSDB; AAB29593.
XX
PT  Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT  diagnosing and treating ion channel disorders, especially long QT
PT  syndrome -
XX
PS  Claim 56; Page -; 132pp; English.
XX
CC  The invention relates to novel ion channel proteins related to
CC  KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC  the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC  respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC  respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC  respectively). The cDNAs encoding these proteins are given in AAC64071-
CC  AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC  potassium channels (I-KR), mutations in which are associated with long
CC  QT syndrome. The invention also relates to methods of diagnosing long QT
```

```
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KCNE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents DNA encoding a mutant human KCNE2
CC   (MiRP1) specifically claimed for use in diagnostic and drug screening
CC   methods of the invention.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the wild-type human KNCE2 cDNA sequence shown on page
CC   118-119.
XX
SQ   Sequence 732 BP; 221 A; 151 C; 158 G; 202 T; 0 other;

Query Match             89.5%;  Score 162;  DB 21;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 2.4e-72;
Matches  162;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          20 AGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTTATTATGGACAATTGGCGCCAGA 79
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          99 AGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTTATTATGGACAATTGGCGCCAGA 158

Qy          80 ACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATG 139
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         159 ACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATG 218

Qy         140 TCATCCTGTACCTCATCATGGTGATGATTGGAATGTTCTCTTTCA 181
               |||||||||||||||||||||||||||||||||||||||||||||
Db         219 TCATCCTGTACCTCATCATGGTGATGATTGGAATGTTCTCTTTCA 260

RESULT 28
AAC64084
ID   AAC64084 standard; DNA; 732 BP.
XX
AC   AAC64084;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Human potassium channel protein KCNE2 (MiRP1) M54T mutant DNA.
XX
KW   Human; KCNE2; MiRP1; potassium channel protein; KNCE1-related;
KW   MinK-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW   fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW   HERG; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200063434-A1.
XX
PD   26-OCT-2000.
XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999; 99US-0129404.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (UYYA ) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672247/65.
DR   P-PSDB; AAB29594.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Claim 56; Page -; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588
CC   respectively; and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KCNE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents DNA encoding a mutant human KCNE2
CC   (MiRP1) specifically claimed for use in diagnostic and drug screening
CC   methods of the invention.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the wild-type human KNCE2 cDNA sequence shown on page
CC   118-119.
XX
SQ   Sequence 732 BP; 221 A; 153 C; 157 G; 201 T; 0 other;

Query Match             85.1%;  Score 154;  DB 21;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 2.7e-68;
Matches  154;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy           1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 139

Qy          61 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         140 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy         121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 154
               ||||||||||||||||||||||||||||||||||
Db         200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 233

RESULT 29
AAD35171
ID   AAD35171 standard; DNA; 732 BP.
XX
AC   AAD35171;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human KCNE2 mutant DNA (T234C).
XX
KW   Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW   KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; mutant; gene; SNP;
KW   single nucleotide polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             74..445
FT                   /*tag= a
FT                   /product= "Human MiRP1 mutant protein"
FT   variation       replace (234, T)
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
XX
```

PN WO200222875-A2.
XX
PD 21-MAR-2002.
XX
PF 11-SEP-2001; 2001WO-US28332.
XX
PR 11-SEP-2000; 2000US-231571P.
XX
PA (UYYA) UNIV YALE.
XX
PI Goldstein SAN;
XX
DR WPI; 2002-362360/39.
DR P-PSDB; AAE22096.
XX
PT Novel gene encoding Min-K related ion channel protein subunit and
PT polymorphisms in this gene associated with antibiotic-induced long QT
PT syndrome -
XX
PS Claim 12; Page 44-45; 49pp; English.
XX
CC The present invention relates to novel KCNE2 genes encoding Min-K related
CC (MiRP) 1 ion channel proteins and polymorphisms in these genes that are
CC associated with ion channel disorders including antibiotic-induced long
CC QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC 57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC encoding the amino acid positions is useful for diagnosing the presence
CC of a polymorphism that causes drug-induced LQTS. The diagnostic methods
CC are useful in the development of new drug therapies which selectively
CC target one or more KCNE2 polymorphisms that are associated with cardiac
CC arrhythmias. The present sequence is human KCNE2 mutant DNA (T234C).
XX
SQ Sequence 732 BP; 221 A; 153 C; 157 G; 201 T; 0 other;

Query Match   85.1%; Score 154; DB 24; Length 732;
Best Local Similarity   100.0%; Pred. No. 2.7e-68;
Matches 154; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 139

Qy   61 ATGGACAATTGCGCGCCAGAACAACACAGCTCAGCAAGAGGCCCTCCAAGCCAAGTTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 ATGGACAATTGCGCGCCAGAACAACACAGCTCAGCAAGAGGCCCTCCAAGCCAAGTTGAT 199

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 154
              ||||||||||||||||||||||||||||||||||
Db   200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 233

RESULT 30
ABA49938
ID ABA49938 standard; DNA; 312 BP.
XX
AC ABA49938;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #8633.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
PR 04-FEB-2000; 2000US-0180312.
XX

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0612366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 4; SEQ ID NO 8633; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   74.0%; Score 134; DB 22; Length 312;
Best Local Similarity   100.0%; Pred. No. 3.6e-58;
Matches 134; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   48 TTTTATTACTTATATGGACAATTGCGCGCCAGAACAACAGCTGAGCAAGCAAGAGGCCCTCCA 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTATTACTTATATGGACAATTGCGCGCCAGAACAACAGCTGAGCAAGCAAGAGGCCCTCCA 60

Qy   108 AGCCAAGTTGATGCTGAGAACTTCTACTATGTCATCGTGATGATTGG 167
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCAAGTTGATGCTGAGAACTTCTACTATGTCATCGTGATGATTGG 120

Qy   168 AATGTTCTCTTTCA 181
              ||||||||||||||
Db   121 AATGTTCTCTTTCA 134

RESULT 31
ABA67856
ID ABA67856 standard; DNA; 312 BP.
XX
AC ABA67856;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #16161.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX

```
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483347/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 16161; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             74.0%;  Score 134;  DB 22;  Length 312;
Best Local Similarity  100.0%;  Pred. No. 3.6e-58;
Matches  134;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     48 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy    108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy    168 AATGTTCTCTTTCA 181
          ||||||||||||||
Db    121 AATGTTCTCTTTCA 134

RESULT 32
ABA34913
ID   ABA34913 standard; DNA; 312 BP.
XX
AC   ABA34913;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Probe #13379 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO20157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
```

```
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 13379; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             74.0%;  Score 134;  DB 22;  Length 312;
Best Local Similarity  100.0%;  Pred. No. 3.6e-58;
Matches  134;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     48 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy    108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy    168 AATGTTCTCTTTCA 181
          ||||||||||||||
Db    121 AATGTTCTCTTTCA 134

RESULT 33
AAK16264
ID   AAK16264 standard; DNA; 312 BP.
XX
AC   AAK16264;
XX
DT   05-NOV-2001 (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16255.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO20157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
```

```
XX
PR    04-FEB-2000; 2000US-0180312.
PR    26-MAY-2000; 2000US-0207456.
PR    30-JUN-2000; 2000US-0608408.
.PR   03-AUG-2000; 2000US-0632366.
PR    21-SEP-2000; 2000US-0234687.
PR    27-SEP-2000; 2000US-0236359.
PR    04-OCT-2000; 2000GB-0024263.
XX
PA    (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI    Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR    WPI; 2001-4B3446/52.
XX
PT    Single exon nucleic acid probes for analyzing gene expression in human
PT    brains -
XX
PS    Example 4; SEQ ID NO: 16255; 650pp + Sequence Listing; English.
XX
CC    The present invention provides a number of single exon nucleic acid
CC    probes which are derived from genomic sequences expressed in the human
CC    brain. They can be used to measure gene expression in brain cell samples,
CC    which may enable the diagnosis and improved treatment of nervous system
CC    diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC    epilepsy and cancers. The present sequence is one of the probes of the
CC    invention.
XX
SQ    Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             74.0%;  Score 134;  DB 22;  Length 312;
Best Local Similarity  100.0%;  Pred. No. 3.6e-58;
Matches  134;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     48 TTTTATTACTTATATGGACAATTGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTATTACTTATATGGACAATTGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy    108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy    168 AATGTTCTCTTTCA 181
          ||||||||||||||
Db    121 AATGTTCTCTTTCA 134

RESULT 34
AAK42008
ID    AAK42008 standard; DNA; 312 BP.
XX
AC    AAK42008;
XX
DT    06-NOV-2001  (first entry)
XX
DE    Human bone marrow expressed single exon probe SEQ ID NO: 16565.
XX
KW    Human; bone marrow expressed exon; gene expression analysis; probe;
KW    microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS    Homo sapiens.
XX
PN    WO200157276-A2.
XX
PD    09-AUG-2001.
XX
PF    30-JAN-2001; 2001WO-US00668.
XX
PR    04-FEB-2000; 2000US-0180312.
PR    26-MAY-2000; 2000US-0207456.
PR    30-JUN-2000; 2000US-0608408.
PR    03-AUG-2000; 2000US-0632366.
PR    21-SEP-2000; 2000US-0234687.
```

```
PR    27-SEP-2000; 2000US-0236359.
PR    04-OCT-2000; 2000GB-0024263.
XX
PA    (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI    Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR    WPI; 2001-488900/53.
XX
PT    Human genome-derived single exon nucleic acid probes useful for
PT    analyzing gene expression in human bone marrow -
XX
PS    Example 4; SEQ ID NO: 16565; 658pp + Sequence Listing; English.
XX
CC    The present invention provides a number of single exon nucleic acid
CC    probes which are derived from genomic sequences expressed in the human
CC    bone marrow. They can be used to measure gene expression in bone marrow
CC    samples, which may enable the improved diagnosis and treatment of cancers
CC    such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC    the probes of the invention.
XX
SQ    Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match             74.0%;  Score 134;  DB 22;  Length 312;
Best Local Similarity  100.0%;  Pred. No. 3.6e-58;
Matches  134;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     48 TTTTATTACTTATATGGACAATTGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTATTACTTATATGGACAATTGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy    108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy    168 AATGTTCTCTTTCA 181
          ||||||||||||||
Db    121 AATGTTCTCTTTCA 134

RESULT 35
AAI22773
ID    AAI22773 standard; DNA; 312 BP.
XX
AC    AAI22773;
XX
DT    12-OCT-2001  (first entry)
XX
DE    Probe #12706 for gene expression analysis in human cervical cell sample.
XX
KW    Probe; human; microarray; gene expression; cervical epithelial cell;
KW    cervical cancer; ss.
XX
OS    Homo sapiens.
XX
PN    WO200157278-A2.
XX
PD    09-AUG-2001.
XX
PF    30-JAN-2001; 2001WO-US00670.
XX
PR    04-FEB-2000; 2000US-0180312.
PR    26-MAY-2000; 2000US-0207456.
PR    30-JUN-2000; 2000US-0608408.
PR    03-AUG-2000; 2000US-0632366.
PR    21-SEP-2000; 2000US-0234687.
PR    27-SEP-2000; 2000US-0236359.
PR    04-OCT-2000; 2000GB-0024263.
XX
PA    (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI    Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
```

```
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 12706; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match      74.0%;   Score 134;  DB 22;  Length 312;
Best Local Similarity   100.0%;  Pred. No. 3.6e-58;
Matches 134;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    48 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy   108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy   168 AATGTTCTCTTTCA 181
         ||||||||||||||
Db   121 AATGTTCTCTTTCA 134

RESULT 36
AAI48075
ID   AAI48075 standard; DNA; 312 BP.
XX
AC   AAI48075;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #16761 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis; ss;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 16761; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match      74.0%;   Score 134;  DB 22;  Length 312;
Best Local Similarity   100.0%;  Pred. No. 3.6e-58;
Matches 134;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    48 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTATTACTTATATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy   108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy   168 AATGTTCTCTTTCA 181
         ||||||||||||||
Db   121 AATGTTCTCTTTCA 134

RESULT 37
AAI08446
ID   AAI08446 standard; DNA; 312 BP.
XX
AC   AAI08446;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #8437 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0612366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 8437; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
```

CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   74.0%; Score 134; DB 22; Length 312;
Best Local Similarity 100.0%; Pred. No. 3.6e-58;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   48 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy   108 AGCCAAAGTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCAAAGTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy   168 AATGTTCTCTTTCA 181
         ||||||||||||||
Db   121 AATGTTCTCTTTCA 134

RESULT 38
ABS16039
ID ABS16039 standard; DNA; 312 BP.
XX
PN WO200186003-A2.
XX
AC ABS16039;
XX
PD 15-NOV-2001.
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 16030.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleimyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -

PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.

XX
PS Claim 4; SEQ ID No 16030; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 312 BP; 94 A; 73 C; 71 G; 74 T; 0 other;

Query Match   74.0%; Score 134; DB 24; Length 312;
Best Local Similarity 100.0%; Pred. No. 3.6e-58;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   48 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy   108 AGCCAAAGTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCAAAGTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy   168 AATGTTCTCTTTCA 181
         ||||||||||||||
Db   121 AATGTTCTCTTTCA 134

RESULT 39
ABN29241
ID ABN29241 standard; DNA; 65 BP.
XX
AC ABN29241;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:1989.
XX

```
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PR   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 1989; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 11 A; 18 C; 15 G; 21 T; 0 other;

Query Match             14.4%;  Score 26;  DB 24;  Length 65;
Best Local Similarity   100.0%;  Pred. No. 0.002;
Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         139 GTCATCCTGTACCTCATGGTGATGAT 164
               ||||||||||||||||||||||||||
Db          13 GTCATCCTGTACCTCATGGTGATGAT 38

RESULT 40
AAS00246
ID   AAS00246 standard; DNA; 372 BP.
XX
AC   AAS00246;
XX
DT   10-MAY-2001   (first entry)
XX
DE   Rat potassium channel regulatory protein, Mink2, DNA sequence.
XX
```

```
KW   Rat; Mink2; potassium channel; cardiac arrhythmia; hypertension; ds;
KW   angina; asthma; diabetes; renal insufficiency; urinary incontinence;
KW   irritable colon; epilepsy; cerebrovascular ischaemia; autoimmune disease.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..372
FT                   /*tag= a
FT                   /product= "MINK2 potassium channel protein"
XX
PN   WO200114403-A1.
XX
PD   01-MAR-2001.
XX
PF   18-AUG-2000; 2000WO-US22799.
XX
PR   20-AUG-1999;  99US-0379201.
XX
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI   Ficker E, Wible B, Brown AM;
XX
DR   WPI; 2001-218424/22.
DR   P-PSDB; AAS00216.
XX
PT   Novel potassium channel gene termed MinK2 encoding potassium channel
PT   regulatory protein, useful for screening compounds that are useful for
PT   treating diseases caused by aberrant potassium activity -
XX
PS   Claim 1; Fig 10; 39pp; English.
XX
CC   The sequence represents the coding sequence of rat potassium channel
CC   regulatory protein, Mink2. Mink2 sequence is useful for producing a
CC   potassium channel regulatory protein useful for in vitro or in vivo
CC   screening of agonistic or antagonistic compounds that are useful for
CC   treating diseases caused by aberrant potassium activity, such as human
CC   cardiac arrhythmias, hypertension, angina, asthma, diabetes, renal
CC   insufficiency, urinary incontinence, irritable colon, epilepsy,
CC   cerebrovascular ischaemia, and autoimmune disease.
XX
SQ   Sequence 372 BP; 95 A; 98 C; 105 G; 74 T; 0 other;

Query Match             14.4%;  Score 26;  DB 22;  Length 372;
Best Local Similarity   100.0%;  Pred. No. 0.0021;
Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         139 GTCATCCTGTACCTCATGGTGATGAT 164
               ||||||||||||||||||||||||||
Db         145 GTCATCCTGTACCTCATGGTGATGAT 170

RESULT 41
AAC64072
ID   AAC64072 standard; cDNA; 468 BP.
XX
AC   AAC64072;
XX
DT   19-FEB-2001   (first entry)
XX
DE   Rat potassium channel protein KCNE2 (MiRP1) cDNA, SEQ ID NO:3.
XX
KW   Rat; KCNE2; MiRP1; potassium channel protein; KNCE1-related;
KW   MinK-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; ion channel disorder;
KW   fast delayed rectifier potassium channel; anti-KNCE2 antibody;
KW   HERG; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200063434-A1.
XX
PD   26-OCT-2000.
```

XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999;   99US-0129404.
XX
PA   (UTAH) UNIV UTAH RES FOUND.
PA   (UYYA) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672747/65.
DR   P-PSDB; AAB29586.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Claim 1; Page 119-120; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC   respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KNCE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents cDNA encoding rat KCNE2 (MiRP1).
XX
SQ   Sequence 468 BP; 118 A; 126 C; 131 G; 93 T; 0 other;

Query Match   14.4%;   Score 26;   DB 21;   Length 468;
Best Local Similarity   100.0%;   Pred. No. 0.0021;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      139 GTCATCCTGTACCTCATGGTGATGAT 164
            ||||||||||||||||||||||||||
Db      179 GTCATCCTGTACCTCATGGTGATGAT 204
```

RESULT 42
AAC64079
ID   AAC64079 standard; DNA; 23 BP.
XX
AC   AAC64079;
XX
DT   19-FEB-2001 (first entry)
XX
DE   PCR primer SEQ ID NO:15, used in KCNE2/3/4 SSCP analysis.
XX
KW   KCNE2; MiRP1; KCNE3; MiRP2; KCNE4; MiRP3; ion channel protein;
KW   KNCE1-related; Mink-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; SSCP analysis;
KW   PCR primer; ss.
XX
OS   Unidentified.
XX
PN   WO200063434-Al.
XX
PD   26-OCT-2000.
XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999;   99US-0129404.
XX
PA   (UTAH) UNIV UTAH RES FOUND.
PA   (UYYA) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672747/65.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Example 1; Page 62; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively); human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC   respectively); and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KNCE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents a PCR primer used to amplify KNCE2,
CC   KNCE3 and KNCE4 gene fragments for SSCP analysis.
XX
SQ   Sequence 23 BP; 5 A; 4 C; 4 G; 10 T; 0 other;

Query Match   12.7%;   Score 23;   DB 21;   Length 23;
Best Local Similarity   100.0%;   Pred. No. 0.065;
Matches   23;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy       34 GTCTTCCGAAGGATTTTATTAC 56
            |||||||||||||||||||||||
Db        1 GTCTTCCGAAGGATTTTATTAC 23
```

RESULT 43
AAC64078/c
ID   AAC64078 standard; DNA; 21 BP.
XX
AC   AAC64078;
XX
DT   19-FEB-2001 (first entry)
XX
DE   PCR primer SEQ ID NO:14, used in KCNE2/3/4 SSCP analysis.
XX
KW   KCNE2; MiRP1; KCNE3; MiRP2; KCNE4; MiRP3; ion channel protein;
KW   KNCE1-related; Mink-related; long QT syndrome; cardiac arrhythmia;
KW   drug screening; knockout mouse; transgenic animal; SSCP analysis;
KW   PCR primer; ss.
XX
OS   Unidentified.
XX
PN   WO200063434-Al.
XX
PD   26-OCT-2000.
XX
PF   14-APR-2000; 2000WO-US10004.
XX
PR   15-APR-1999;   99US-0129404.

```
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (UYYA ) UNIV YALE.
XX
PI   Abbott GW, Sesti F, Splawski I, Keating MT, Goldstein SAN;
XX
DR   WPI; 2000-672747/65.
XX
PT   Novel nucleic acids encoding MiRP1, MiRP2 and MiRP3, useful for
PT   diagnosing and treating ion channel disorders, especially long QT
PT   syndrome -
XX
PS   Example 1; Page 62; 132pp; English.
XX
CC   The invention relates to novel ion channel proteins related to
CC   KNCE1 (MinK) and to nucleic acids encoding them. The proteins of
CC   the invention are human and rat KCNE2 (MiRP1; AAB29585 and AAB29586,
CC   respectively; human and mouse KCNE3 (MiRP2; AAB29587 and AAB29588,
CC   respectively; and human and mouse KCNE4 (MiRP3; AAB29589 and AAB29590,
CC   respectively). The cDNAs encoding these proteins are given in AAC64071-
CC   AAC64076. KCNE2, along with HERG, forms cardiac fast delayed rectifier
CC   potassium channels (I-KR), mutations in which are associated with long
CC   QT syndrome. The invention also relates to methods of diagnosing long QT
CC   syndrome using the KCNE2, KCNE3 or KCNE4 genes, a knockout mouse with a
CC   disruption in an endogenous KCNE2, KCNE3 or KCNE4 gene, transgenic
CC   nonhuman animals comprising a heterologous ion channel protein gene
CC   of the invention, a transgenic animal comprising human KNCE2 and HERG
CC   DNA, and methods of and screening drugs for treating long QT syndrome
CC   using KCNE2 proteins (including mutants), nucleic acids encoding them
CC   and antibodies against KNCE2 proteins. The methods, antibodies, nucleic
CC   acids, and proteins may be used for diagnosing or treating ion channel
CC   disorders, especially long QT syndrome. Transgenic animals comprising
CC   KCNE2 and HERG are useful for testing anti-long QT syndrome drugs.
CC   The present sequence represents a PCR primer used to amplify KNCE2,
CC   KNCE3 and KNCE4 gene fragments for SSCP analysis.
XX
SQ   Sequence 21 BP; 5 A; 4 C; 6 G; 6 T; 0 other;

Query Match           11.6%;  Score 21;  DB 21;  Length 21;
Best Local Similarity  100.0%;  Pred. No. 0.67;
Matches   21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   103 CTCCAAGCCAAAGTTGATGCT 123
         |||||||||||||||||||||
Db    21 CTCCAAGCCAAAGTTGATGCT 1

RESULT 44
AAD35176
ID   AAD35176 standard; DNA; 21 BP.
XX
AC   AAD35176;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human KCNE2 gene amplifying forward PCR primer #2.
XX
KW   Human; Min-K related ion channel protein; MiRP1; ion channel disorder;
KW   KCNE2; long QT syndrome; LQTS; cardiac arrhythmia; PCR; primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200222875-A2.
XX
PD   21-MAR-2002.
XX
PF   11-SEP-2001; 2001WO-US28332.
XX
PR   11-SEP-2000; 2000US-231571P.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Goldstein SAN;
```

```
XX
DR   WPI; 2002-362360/39.
XX
PT   Novel gene encoding Min-K related ion channel protein subunit and
PT   polymorphisms in this gene associated with antibiotic-induced long QT
PT   syndrome -
XX
PS   Example 1; Page 22; 49pp; English.
XX
CC   The present invention relates to novel KCNE2 genes encoding Min-K related
CC   (MiRP) 1 ion channel proteins and polymorphisms in these genes that are
CC   associated with ion channel disorders including antibiotic-induced long
CC   QT syndrome (LQTS). Detecting a mutation at amino acid positions 8, 54,
CC   57 or 116 of MiRP1 polypeptide or a mutation at a nucleotide position
CC   encoding the amino acid positions is useful for diagnosing the presence
CC   of a polymorphism that causes drug-induced LQTS. The diagnostic methods
CC   are useful in the development of new drug therapies which selectively
CC   target one or more KCNE2 polymorphisms that are associated with cardiac
CC   arrhythmias. The present sequence is human KCNE2 gene amplifying PCR
CC   primer. This sequence is used in the exemplification of the invention.
XX
SQ   Sequence 21 BP; 4 A; 7 C; 2 G; 8 T; 0 other;

Query Match           11.6%;  Score 21;  DB 24;  Length 21;
Best Local Similarity  100.0%;  Pred. No. 0.67;
Matches   21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   133 TACTATGTCATCCTGTACCTC 153
         |||||||||||||||||||||
Db     1 TACTATGTCATCCTGTACCTC 21

RESULT 45
AAF80272
ID   AAF80272 standard; DNA; 20 BP.
XX
AC   AAF80272;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Primer used to amplify potassium channel subunit IsK2 cDNA fragment.
XX
KW   Human; potassium channel; IsK2; gene therapy; gastric motility;
KW   gastric acid secretion; anti-arrhythmic agent; myocardial infarction;
KW   PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200127246-A1.
XX
PD   19-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US28014.
XX
PR   12-OCT-1999; 99US-0158781.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Swanson RJ,  Liu Y,  Folander K;
XX
DR   WPI; 2001-273764/28.
XX
PS   Example 2; Page 31; 46pp; English.
XX
CC   New DNA encoding the IsK2 potassium channel subunit, useful e.g. for
CC   detecting mutations and screening for therapeutic agents -
CC
CC   PCR primers AAF80272-73 were used to amplify a cDNA fragment encoding a
CC   potassium channel subunit, designated IsK2. The IsK2 polynucleotide,
CC   and derived probes, are used diagnostically to detect mutations in the
CC   IsK2 gene, to determine levels of mRNA expression and to isolate
CC   homologous sequences; for recombinant expression of IsK2; in gene
CC   therapy to increase potassium channel activity and to generate
```

```
CC    transgenic animals, as models and for drug screening. Recombinant IsK2
CC    is used for studying biochemical activity of IsK2 and its role in
CC    disorders of gastric motility and gastric acid secretion, and to raise
CC    specific antibodies. IsK2 modulators are potentially useful for
CC    treating diseases associated with increased or reduced potassium
CC    channel activity, e.g. as anti-arrhythmic agents for treating
CC    myocardial infarction and as regulators of gastric acid secretion.
XX
SQ    Sequence 20 BP; 5 A; 3 C; 6 G; 6 T; 0 other;

Query Match            11.0%;  Score 20;  DB 22;  Length 20;
Best Local Similarity 100.0%;  Pred. No. 2.2;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      150 CCTCATGGTGATGATTGGAA 169
            ||||||||||||||||||||
Db        1 CCTCATGGTGATGATTGGAA 20


Search completed: June  9, 2003, 12:12:24
Job time : 184 secs
```

```
                        GenCore version 5.1.6
            Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     June  9, 2003, 12:33:26 ; Search time 82 Seconds
                                (without alignments)
                                3089.223 Million cell updates/sec

Title:         US-09-550-163-1_COPY_80_260
Perfect score: 181
Sequence:      1 acttatccaattcacaca.........gattgaatgtctctttca 181

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched:      870385 seqs, 699768693 residues

Word size :    0

Total number of hits satisfying chosen parameters:      1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :     Published Applications NA:*
               1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
               2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
               3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
               4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
               5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
               6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
               7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
               8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
               9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
              10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
              11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
              12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
              13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
              14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	c	Score	% Match	Query Length	DB	ID	Description
1		181	100.0	372	10	US-09-864-761-33139	Sequence 33139, A
2		181	100.0	450	10	US-09-864-761-3463	Sequence 3463, Ap
3		181	100.0	471	10	US-09-864-761-16671	Sequence 16671, A
4		181	100.0	732	9	US-10-000-151B-5	Sequence 5, Appli
5		181	100.0	113604	9	US-10-227-195A-1	Sequence 1, Appli
6		181	100.0	113604	9	US-10-227-195A-2	Sequence 2, Appli
7		134	74.0	312	10	US-09-864-761-20233	Sequence 20233, A
8		18	9.9	714	10	US-09-867-701-10563	Sequence 10563, A
9	c	18	9.9	782	10	US-09-772-134B-89	Sequence 89, Appl
10		17	9.4	374	10	US-09-867-701-7019	Sequence 7019, Appl
11		17	9.4	375	10	US-09-867-701-7022	Sequence 7022, Ap
12		17	9.4	475	9	US-09-918-995-69	Sequence 69, Appli
13		17	9.4	491	9	US-09-918-995-1925	Sequence 1925, Ap
14	c	17	9.4	2197	10	US-09-778-171-1	Sequence 1, Appli
15		17	9.4	32189	9	US-09-764-891-7358	Sequence 7358, Ap
16	c	17	9.4	75899	10	US-09-854-883-243	Sequence 243, Appli
17	c	17	9.4	465237	9	US-09-933-267A-1	Sequence 1, Appli
18	c	17	9.4	1691139	9	US-10-067-514-1	Sequence 1, Appli
19	c	16	8.8	324	10	US-09-960-352-12908	Sequence 12908, A
20	c	16	8.8	438	10	US-09-983-965-2960	Sequence 2960, Ap
21		16	8.8	622	10	US-09-879-536-381	Sequence 381, App
22		16	8.8	1977	10	US-09-938-842A-1502	Sequence 1502, Ap
23		16	8.8	2244	10	US-09-954-455-1174	Sequence 1174, Ap
24		16	8.8	2244	10	US-09-954-456-1838	Sequence 1838, App
25		16	8.8	2880	10	US-09-925-301-415	Sequence 415, Appl
26		16	8.8	2889	10	US-09-851-129A-16	Sequence 16, Appl
27	c	16	8.8	6047	10	US-09-954-456-498	Sequence 498, App
28		16	8.8	11990	10	US-09-969-708-569	Sequence 569, App
29		16	8.8	14796	9	US-10-138-618-35	Sequence 35, Appl
30		16	8.8	14796	10	US-09-954-456-973	Sequence 973, App
31		16	8.8	14796	10	US-09-954-456-1636	Sequence 1636, App
32		16	8.8	14796	10	US-09-918-186A-3	Sequence 3, Appli
33		16	8.8	14796	10	US-09-880-107-3421	Sequence 3421, Ap
34		16	8.8	31314	10	US-09-764-877-3875	Sequence 3875, App
35		16	8.8	123192	9	US-10-175-523-71	Sequence 71, Appl
36		16	8.8	186957	9	US-10-185-770-3	Sequence 3, Appli
37	c	16	8.8	640681	9	US-09-790-988-1	Sequence 1, Appli
38	c	16	8.8	684973	9	US-09-263-959-1	Sequence 1, Appli
39		15	8.3	81	10	US-09-864-761-23326	Sequence 23326, A
40		15	8.3	133	10	US-09-560-863-157	Sequence 157, App
41		15	8.3	136	10	US-09-770-696-583	Sequence 583, App
42	c	15	8.3	246	10	US-09-991-936-464	Sequence 464, App
43		15	8.3	271	10	US-09-864-761-20062	Sequence 20062, A
44		15	8.3	271	10	US-09-878-574-11763	Sequence 11763, A
45		15	8.3	274	10	US-09-864-761-19038	Sequence 19038, A
46		15	8.3	274	10	US-09-864-761-19234	Sequence 19234, A
47	c	15	8.3	280	10	US-09-783-590-2236	Sequence 2236, Ap
48	c	15	8.3	285	10	US-09-924-035A-336	Sequence 336, App
49		15	8.3	295	10	US-09-294-093B-748	Sequence 748, App
50		15	8.3	338	9	US-09-918-995-18704	Sequence 18704, A
51		15	8.3	346	9	US-09-954-531-521	Sequence 521, App
52		15	8.3	348	9	US-09-918-995-18030	Sequence 18030, A
53		15	8.3	369	9	US-09-803-719-1377	Sequence 1377, Ap
54		15	8.3	372	10	US-09-728-446-1042	Sequence 1042, Ap
55	c	15	8.3	412	9	US-10-060-036-293	Sequence 293, App
56		15	8.3	420	9	US-09-902-180-5	Sequence 5, Appli
57		15	8.3	424	10	US-09-918-995-35140	Sequence 35140, A
58		15	8.3	451	9	US-09-918-995-3384	Sequence 3384, Ap
59		15	8.3	453	9	US-09-764-891-9895	Sequence 9895, Ap
60	c	15	8.3	455	10	US-09-864-761-2303	Sequence 2303, Ap
61	c	15	8.3	462	9	US-09-918-995-14296	Sequence 14296, A
62	c	15	8.3	468	9	US-09-918-995-9640	Sequence 9640, Ap
63		15	8.3	472	10	US-09-864-761-3287	Sequence 3287, Ap
64		15	8.3	475	10	US-09-864-761-2506	Sequence 2506, Ap
65		15	8.3	501	10	US-09-783-590-4657	Sequence 4657, Ap
66		15	8.3	516	9	US-09-738-626-564	Sequence 564, App
67		15	8.3	541	10	US-09-864-761-8318	Sequence 8318, Ap
68		15	8.3	544	9	US-09-918-995-26895	Sequence 26895, A
69		15	8.3	555	10	US-09-864-761-8868	Sequence 8868, Ap
70		15	8.3	563	9	US-10-102-806-11	Sequence 11, Appl
71		15	8.3	569	9	US-09-796-692-8178	Sequence 8178, Ap
72		15	8.3	569	10	US-10-040-862-8178	Sequence 8178, Ap
73		15	8.3	593	10	US-09-910-943-260	Sequence 260, App
74	c	15	8.3	701	10	US-09-864-761-6592	Sequence 6592, Ap
75		15	8.3	714	9	US-10-198-846-7544	Sequence 7544, Ap
76		15	8.3	833	9	US-10-198-846-7267	Sequence 7267, Ap
77		15	8.3	842	9	US-10-198-846-80	Sequence 80, Appl
78		15	8.3	858	9	US-10-198-846-3297	Sequence 3297, Ap
79		15	8.3	861	9	US-10-198-846-1685	Sequence 1685, Ap
80		15	8.3	871	9	US-10-198-846-1728	Sequence 1728, Ap
81		15	8.3	882	9	US-10-198-846-3314	Sequence 3314, Ap
82		15	8.3	896	9	US-10-198-846-997	Sequence 997, Appl
83		15	8.3	1135	9	US-09-822-846-16	Sequence 16, Appl
84		15	8.3	1225	9	US-09-822-846-198	Sequence 198, App
85		15	8.3	1240	9	US-09-847-101B-28	Sequence 28, Appl
86	c	15	8.3	1251	10	US-09-925-300-684	Sequence 684, App
87		15	8.3	1385	12	US-10-044-090-700	Sequence 700, App
88		15	8.3	1396	9	US-09-764-864-220	Sequence 220, App
89	c	15	8.3	1425	10	US-10-102-806-2	Sequence 2, Appli
90		15	8.3	1500	10	US-09-922-501-9	Sequence 9, Appli
91	c	15	8.3	2446	10	US-09-098-807-1	Sequence 1, Appli
92	c	15	8.3	2535	10	US-09-815-242-4378	Sequence 4378, Ap

	idx				US reference		Sequence
c	93	15	8.3	2610	US-09-815-242-8168	10	Sequence 8168, Ap
c	94	15	8.3	2682	US-09-880-107-3300	10	Sequence 3300, Ap
c	95	15	8.3	3261	US-09-005-216-1	10	Sequence 1, Appli
c	96	15	8.3	3353	US-09-764-864-96	10	Sequence 96, Appli
	97	15	8.3	3842	US-09-978-729A-24	9	Sequence 24, Appl
	98	15	8.3	3842	US-09-981-087A-24	9	Sequence 24, Appl
	99	15	8.3	3842	US-09-978-382A-24	9	Sequence 24, Appl
	100	15	8.3	3842	US-09-978-740A-24	9	Sequence 24, Appl
	101	15	8.3	3997	US-09-978-730-24	10	Sequence 223, Appp
	102	15	8.3	3997	US-08-781-986A-223	7	Sequence 223, Appp
	103	15	8.3	5428	US-09-533-029-119	9	Sequence 119, App
	104	15	8.3	5774	US-10-239-676-193	9	Sequence 193, App
c	105	15	8.3	6022	US-10-239-676-80	9	Sequence 80, Appl
	106	15	8.3	7231	US-09-847-101B-42	9	Sequence 42, Appl
c	107	15	8.3	7601	US-09-764-877-2897	9	Sequence 2897, Ap
c	108	15	8.3	7601	US-09-764-877-2898	10	Sequence 2898, App
c	109	15	8.3	7960	US-09-847-101B-30	9	Sequence 30, Appl
	110	15	8.3	7989	US-09-847-101B-33	9	Sequence 33, Appl
	111	15	8.3	8383	US-09-847-101B-29	9	Sequence 29, Appl
	112	15	8.3	11960	US-09-764-891-8501	9	Sequence 8501, Ap
c	113	15	8.3	11991	US-09-764-891-9893	9	Sequence 9893, Ap
	114	15	8.3	11991	US-09-764-877-2942	10	Sequence 2942, Ap
c	115	15	8.3	14176	US-09-764-864-1644	10	Sequence 1644, Ap
	116	15	8.3	21129	US-10-091-504-1734	10	Sequence 1734, Ap
c	117	15	8.3	21129	US-09-764-869-1734	10	Sequence 1734, Ap
	118	15	8.3	23071	US-09-764-864-1673	10	Sequence 1673, Ap
c	119	15	8.3	32480	US-09-847-101B-23	9	Sequence 23, Appl
	120	15	8.3	33023	US-09-880-107-3350	10	Sequence 3350, Ap
	121	15	8.3	34427	US-09-111-911-5	9	Sequence 5, Appli
	122	15	8.3	35408	US-10-155-649-3	9	Sequence 3, Appli
	123	15	8.3	35871	US-09-956-135-2	9	Sequence 2, Appli
	124	15	8.3	35935	US-09-725-720-43	10	Sequence 43, Appli
	125	15	8.3	35935	US-09-782-178A-4	10	Sequence 4, Appli
	126	15	8.3	35935	US-09-782-178A-5	10	Sequence 5, Appli
	127	15	8.3	35937	US-09-782-178A-3	10	Sequence 3, Appli
	128	15	8.3	35978	US-09-952-060-30	10	Sequence 30, Appl
	129	15	8.3	36620	US-09-952-060-25	10	Sequence 25, Appl
	130	15	8.3	37474	US-09-952-060-28	9	Sequence 28, Appl
	131	15	8.3	38519	US-09-913-514-1	9	Sequence 1, Appl
	132	15	8.3	124884	US-09-913-514-2	10	Sequence 2, Appli
	133	15	8.3	125157	US-09-913-514-17	9	Sequence 17, Appl
	134	15	8.3	1127762	US-09-954-556-17	9	Sequence 62, Appl
	135	15	8.3	185548	US-10-175-523-62	9	Sequence 3, Appli
	136	15	8.3	202001	US-10-274-990-3	10	Sequence 30, Appl
	137	15	8.3	202001	US-09-734-674-3	9	Sequence 25, Appl
	138	15	8.3	368004	US-08-781-986A-2113	7	Sequence 2113, Ap
c	139	15	8.3	378161	US-09-901-136-3	9	Sequence 3, Appli
	140	15	8.3	397658	US-09-813-120-3	10	Sequence 3, Appli
	141	15	8.3	1691139	US-10-067-514-1	10	Sequence 1, Appli
	142	15	8.3	3109400	US-09-738-626-1	8	Sequence 1, Appli
c	143	14	7.7	23	US-09-302-117-1	10	Sequence 120, App
	144	14	7.7	66	US-09-747-377-120	9	Sequence 120, App
c	145	14	7.7	66	US-10-105-613-120	10	Sequence 120, App
	146	14	7.7	87	US-08-781-986A-2113	7	Sequence 2113, Ap
	147	14	7.7	96	US-10-080-797-14	10	Sequence 14, Appl
c	148	14	7.7	125	US-09-783-590-704	10	Sequence 704, Appp
	149	14	7.7	133	US-09-764-761-19101	10	Sequence 19101, A
	150	14	7.7	159	US-09-878-574-13246	10	Sequence 13246, Ap
	151	14	7.7	162	US-09-828-523A-97	9	Sequence 97, Appl
	152	14	7.7	162	US-09-966-521-89	9	Sequence 89, Appl
c	153	14	7.7	171	US-09-864-761-20817	10	Sequence 20817, A
	154	14	7.7	186	US-09-864-761-20227	10	Sequence 20227, A
	155	14	7.7	195	US-09-960-152-515	10	Sequence 515, App
	156	14	7.7	232	US-09-960-152-9112	9	Sequence 9112, Ap
	157	14	7.7	250	US-09-837-459-25	10	Sequence 25, Appl
	158	14	7.7	251	US-08-837-574-5617	10	Sequence 5617, App
c	159	14	7.7	253	US-09-867-701-10436	10	Sequence 10436, A
	160	14	7.7	258	US-09-878-574-9246	10	Sequence 9246, Ap
	161	14	7.7	259	US-09-878-598-2369	9	Sequence 2369, Ap
c	162	14	7.7	266	US-09-815-242-3312	10	Sequence 3312, Ap
	163	14	7.7	284	US-09-783-590-3447	10	Sequence 3447, Ap
c	164	14	7.7	289	US-09-764-877-782	10	Sequence 782, Appp
c	165	14	7.7	292	US-09-294-093B-546	10	Sequence 546, App

	idx					US reference		Sequence
	166	14	7.7	300	10	US-09-294-093B-627		Sequence 627, App
c	167	14	7.7	301	10	US-09-563-817-238		Sequence 238, App
	168	14	7.7	302	9	US-09-796-692-3202		Sequence 3202, Ap
	169	14	7.7	302	9	US-10-040-862-3202		Sequence 3202, Ap
c	170	14	7.7	310	9	US-10-060-036-1631		Sequence 1631, Ap
	171	14	7.7	321	10	US-09-960-352-10556		Sequence 10556, A
c	172	14	7.7	324	10	US-09-867-550-551		Sequence 551, App
	173	14	7.7	329	10	US-09-770-791-993		Sequence 993, App
	174	14	7.7	330	10	US-09-864-761-32635		Sequence 32635, A
	175	14	7.7	335	10	US-09-934-586A-12		Sequence 12, Appl
	176	14	7.7	344	10	US-09-983-965-2895		Sequence 2895, Ap
c	177	14	7.7	364	9	US-09-918-995-37558		Sequence 37558, A
	178	14	7.7	364	10	US-09-878-574-1638		Sequence 1638, Ap
c	179	14	7.7	367	9	US-09-918-995-37588		Sequence 37588, A
	180	14	7.7	368	9	US-09-918-995-19075		Sequence 19075, A
c	181	14	7.7	372	10	US-09-864-761-33139		Sequence 33139, A
c	182	14	7.7	375	9	US-10-066-543-2800		Sequence 2800, Ap
c	183	14	7.7	381	9	US-09-983-965-3886		Sequence 3886, Ap
	184	14	7.7	382	10	US-09-983-965-2099		Sequence 2099, Ap
c	185	14	7.7	389	10	US-09-960-352-5803		Sequence 5803, Ap
	186	14	7.7	391	10	US-09-728-446-528		Sequence 528, App
	187	14	7.7	391	10	US-09-880-107-816		Sequence 816, Appp
c	188	14	7.7	393	9	US-10-152-661-26		Sequence 26, Appl
c	189	14	7.7	393	9	US-09-866-050A-26		Sequence 26, Appl
	190	14	7.7	395	10	US-09-960-352-12338		Sequence 12338, A
	191	14	7.7	397	9	US-09-983-965-1897		Sequence 1897, Ap
	192	14	7.7	399	9	US-09-918-995-8158		Sequence 8158, Ap
	193	14	7.7	402	10	US-09-918-995-15232		Sequence 15232, A
	194	14	7.7	411	9	US-09-864-761-3457		Sequence 3457, Ap
	195	14	7.7	412	9	US-09-946-807-1419		Sequence 1419, Ap
c	196	14	7.7	412	9	US-10-198-846-947		Sequence 947, Appp
	197	14	7.7	412	10	US-09-795-668-1419		Sequence 1419, Ap
	198	14	7.7	412	10	US-09-795-686-1419		Sequence 1419, Ap
	199	14	7.7	413	10	US-09-864-761-4499		Sequence 4499, Ap
c	200	14	7.7	415	10	US-09-864-761-4061		Sequence 4061, Ap
c	201	14	7.7	420	10	US-09-918-995-35144		Sequence 35144, A
	202	14	7.7	421	10	US-09-960-352-10402		Sequence 10402, A
	203	14	7.7	424	10	US-09-960-352-1175		Sequence 1175, Ap
	204	14	7.7	428	10	US-09-983-965-3430		Sequence 3430, Ap
	205	14	7.7	430	10	US-09-983-965-4101		Sequence 4101, Ap
	206	14	7.7	433	10	US-09-815-341-1123		Sequence 1123, Ap
	207	14	7.7	434	10	US-09-919-580-133		Sequence 133, App
c	208	14	7.7	441	10	US-10-046-935-176		Sequence 176, App
c	209	14	7.7	441	10	US-09-878-178-176		Sequence 176, App
c	210	14	7.7	441	10	US-10-146-502-176		Sequence 176, App
	211	14	7.7	442	10	US-09-983-965-2204		Sequence 2204, Ap
	212	14	7.7	442	9	US-09-796-692-4830		Sequence 4830, Ap
	213	14	7.7	447	9	US-10-040-862-4830		Sequence 4830, Ap
	214	14	7.7	448	9	US-10-198-846-5995		Sequence 5995, Ap
	215	14	7.7	450	10	US-09-962-436-548		Sequence 548, Appp
	216	14	7.7	450	10	US-09-864-761-3463		Sequence 3463, Ap
	217	14	7.7	456	10	US-09-864-761-6463		Sequence 6463, Ap
	218	14	7.7	457	10	US-09-864-761-2367		Sequence 2367, Ap
c	219	14	7.7	457	9	US-10-198-846-1909		Sequence 1909, Ap
c	220	14	7.7	457	9	US-10-198-846-6134		Sequence 6134, Ap
	221	14	7.7	457	9	US-09-967-768A-42		Sequence 42, Appl
	222	14	7.7	458	9	US-09-918-995-25998		Sequence 25998, A
	223	14	7.7	458	9	US-09-764-891-8273		Sequence 8273, Ap
c	224	14	7.7	459	9	US-09-736-457-47		Sequence 47, Appl
c	225	14	7.7	459	9	US-09-902-941-47		Sequence 47, Appl
c	226	14	7.7	459	9	US-09-849-626-47		Sequence 47, Appl
c	227	14	7.7	459	9	US-10-017-754-47		Sequence 47, Appl
c	228	14	7.7	461	9	US-09-918-995-31584		Sequence 31584, A
	229	14	7.7	462	9	US-10-046-935-28		Sequence 28, Appl
	230	14	7.7	462	9	US-09-878-178-28		Sequence 28, Appl
	231	14	7.7	462	9	US-10-146-502-28		Sequence 28, Appl
	232	14	7.7	465	10	US-09-816-523-4		Sequence 4, Appli
c	233	14	7.7	470	10	US-10-091-504-1933		Sequence 1933, Ap
	234	14	7.7	470	10	US-09-764-869-1933		Sequence 1933, Ap
	235	14	7.7	471	9	US-09-918-995-8463		Sequence 8463, Ap
c	236	14	7.7	471	10	US-09-864-761-16671		Sequence 16671, A
c	237	14	7.7	471	10	US-09-864-761-24760		Sequence 24760, A
c	238	14	7.7	472	9	US-09-918-995-15314		Sequence 15314, A

c	#			n	len	US ref	Sequence
c	239	14	7.7	472	9	US-09-918-995-21724	Sequence 21724, A
	240	14	7.7	474	10	US-09-864-761-10497	Sequence 10497, A
c	241	14	7.7	475	9	US-09-918-995-11543	Sequence 11543, A
c	242	14	7.7	475	9	US-09-918-995-32842	Sequence 32842, A
c	243	14	7.7	477	9	US-09-918-995-24664	Sequence 24664, A
	244	14	7.7	478	9	US-09-918-995-20685	Sequence 20685, A
	245	14	7.7	478	10	US-09-918-995-24301	Sequence 24301, A
	246	14	7.7	478	9	US-09-864-761-11151	Sequence 11151, A
c	247	14	7.7	480	9	US-09-764-868-1449	Sequence 1449, A
	248	14	7.7	483	9	US-09-764-868-359	Sequence 359, App
	249	14	7.7	483	10	US-09-833-790-251	Sequence 251, App
	250	14	7.7	483	9	US-09-833-790-406	Sequence 406, App
	251	14	7.7	484	9	US-09-983-965-3773	Sequence 3773, Ap
	252	14	7.7	485	10	US-09-918-995-16274	Sequence 16274, A
	253	14	7.7	488	9	US-09-918-995-10341	Sequence 10341, A
c	254	14	7.7	490	9	US-10-198-846-1261	Sequence 1261, Ap
	255	14	7.7	491	9	US-09-918-995-31313	Sequence 31313, A
c	256	14	7.7	493	9	US-09-918-995-36747	Sequence 36747, A
	257	14	7.7	493	9	US-09-918-995-23723	Sequence 23723, A
c	258	14	7.7	495	10	US-09-878-574-4840	Sequence 4840, Ap
	259	14	7.7	501	9	US-10-198-846-209	Sequence 209, App
	260	14	7.7	504	9	US-09-783-590-8541	Sequence 8541, Ap
	261	14	7.7	505	9	US-09-918-995-4085	Sequence 4085, Ap
c	262	14	7.7	507	9	US-09-918-995-21340	Sequence 21340, A
c	263	14	7.7	523	9	US-10-198-846-5	Sequence 5, Appli
	264	14	7.7	523	9	US-09-764-891-7784	Sequence 7784, Ap
	265	14	7.7	524	10	US-09-998-598-280	Sequence 280, App
	266	14	7.7	527	9	US-09-998-598-535	Sequence 535, App
	267	14	7.7	527	9	US-10-198-846-136	Sequence 136, App
c	268	14	7.7	529	10	US-09-920-300A-110	Sequence 110, App
	269	14	7.7	529	10	US-10-033-528-110	Sequence 110, App
	270	14	7.7	530	9	US-09-864-761-16158	Sequence 16158, A
c	271	14	7.7	542	10	US-09-864-761-23176	Sequence 23176, A
	272	14	7.7	547	9	US-09-917-800A-75	Sequence 75, Appl
	273	14	7.7	547	9	US-09-870-759-131	Sequence 131, App
	274	14	7.7	547	9	US-09-736-457-91	Sequence 91, Appl
	275	14	7.7	549	9	US-09-902-941-91	Sequence 91, Appl
	276	14	7.7	552	9	US-09-849-626-91	Sequence 91, Appl
	277	14	7.7	555	9	US-10-017-754-91	Sequence 91, Appl
	278	14	7.7	560	9	US-09-991-936-1721	Sequence 1721, Ap
c	279	14	7.7	570	9	US-09-864-761-7210	Sequence 7210, Ap
c	280	14	7.7	573	10	US-10-198-846-33	Sequence 33, Appl
	281	14	7.7	575	10	US-09-867-550-1179	Sequence 1179, Ap
	282	14	7.7	578	10	US-09-864-761-9618	Sequence 9618, Ap
	283	14	7.7	578	9	US-09-864-761-16225	Sequence 16225, A
	284	14	7.7	578	9	US-10-198-846-4190	Sequence 4190, Ap
	285	14	7.7	580	9	US-10-025-380-422	Sequence 422, App
	286	14	7.7	583	10	US-09-922-217-422	Sequence 422, App
	287	14	7.7	591	10	US-09-833-263-422	Sequence 422, App
	288	14	7.7	596	10	US-09-833-790-108	Sequence 108, App
c	289	14	7.7	599	10	US-09-864-761-8538	Sequence 8538, Ap
c	290	14	7.7	603	10	US-09-924-035A-865	Sequence 865, App
	291	14	7.7	608	9	US-09-864-761-8757	Sequence 8757, Ap
c	292	14	7.7	610	9	US-09-864-761-8229	Sequence 8229, Ap
	293	14	7.7	615	9	US-10-066-543-251	Sequence 251, App
c	294	14	7.7	623	9	US-10-198-846-3837	Sequence 3837, Ap
	295	14	7.7	625	9	US-09-938-842A-3718	Sequence 3718, Ap
	296	14	7.7	625	9	US-09-917-800A-533	Sequence 533, App
	297	14	7.7	626	9	US-10-198-846-16	Sequence 16, Appl
c	298	14	7.7	627	9	US-10-198-846-6978	Sequence 6978, Ap
	299	14	7.7	630	9	US-10-078-090-94	Sequence 94, Appl
	300	14	7.7	646	9	US-10-198-846-6324	Sequence 6324, Ap
	301	14	7.7	647	9	US-10-198-846-3625	Sequence 3625, Ap
	302	14	7.7	650	10	US-09-770-149-728	Sequence 728, App
	303	14	7.7	660	9	US-10-198-846-3620	Sequence 3620, Ap
	304	14	7.7	665	9	US-10-198-846-2053	Sequence 2053, Ap
	305	14	7.7	667	9	US-10-198-846-8281	Sequence 8281, Ap
c	306	14	7.7	667	9	US-10-198-846-3248	Sequence 3248, Ap
c	307	14	7.7	669	9	US-10-198-846-2390	Sequence 2390, Ap
	308	14	7.7		9	US-10-198-846-8441	Sequence 8441, Ap
	309	14	7.7		9	US-10-198-846-7369	Sequence 7369, Ap
	310	14	7.7		10	US-10-198-846-7715	Sequence 7715, Ap
c	311	14	7.7		9	US-10-198-846-2083	Sequence 2083, Ap
c	312	14	7.7	693	9	US-10-198-846-8782	Sequence 8782, Ap
c	313	14	7.7	694	9	US-10-198-846-5280	Sequence 5280, Ap
	314	14	7.7	694	10	US-09-878-574-4825	Sequence 4825, Ap
	315	14	7.7	700	9	US-09-832-129-26	Sequence 26, Appl
	316	14	7.7	700	9	US-10-050-704-96	Sequence 96, Appl
c	317	14	7.7	714	10	US-10-198-846-3906	Sequence 3906, Ap
c	318	14	7.7	721	9	US-09-878-574-4649	Sequence 4649, Ap
	319	14	7.7	726	9	US-10-198-846-9289	Sequence 9289, Ap
	320	14	7.7	727	10	US-10-198-846-10726	Sequence 10726, A
	321	14	7.7	729	10	US-09-878-574-4642	Sequence 4642, Ap
	322	14	7.7	732	10	US-09-919-580-882	Sequence 882, App
c	323	14	7.7	749	10	US-10-000-151B5	Sequence 5, Appli
	324	14	7.7	754	10	US-09-770-149-14	Sequence 14, Appl
	325	14	7.7	754	9	US-09-764-891-8588	Sequence 8588, Ap
	326	14	7.7	755	10	US-09-764-891-8590	Sequence 8590, Ap
	327	14	7.7	759	9	US-09-820-089A-4	Sequence 4, Appli
	328	14	7.7	779	9	US-09-738-626-2187	Sequence 2187, Ap
	329	14	7.7	780	9	US-10-202-193-139	Sequence 139, App
	330	14	7.7	802	9	US-10-198-846-4573	Sequence 4573, Ap
	331	14	7.7	804	9	US-10-198-846-1857	Sequence 1857, Ap
	332	14	7.7	804	10	US-10-198-846-4433	Sequence 443, Ap
	333	14	7.7	810	10	US-09-815-242-9488	Sequence 9488, Ap
	334	14	7.7	810	9	US-10-078-090-95	Sequence 95, Appl
	335	14	7.7	811	9	US-10-198-846-5909	Sequence 5909, Ap
	336	14	7.7	811	9	US-10-202-193-119	Sequence 119, App
	337	14	7.7	814	9	US-10-066-543-68	Sequence 68, Appl
	338	14	7.7	819	9	US-10-071-766-18	Sequence 18, Appl
	339	14	7.7	830	9	US-10-198-846-5835	Sequence 5835, Ap
	340	14	7.7	831	9	US-10-202-193-253	Sequence 253, App
	341	14	7.7	831	10	US-09-815-242-9020	Sequence 9020, Ap
	342	14	7.7	833	9	US-10-198-846-4833	Sequence 4833, Ap
	343	14	7.7	836	9	US-09-764-891-1739	Sequence 1739, Ap
	344	14	7.7	837	9	US-10-202-193-199	Sequence 199, App
	345	14	7.7	838	9	US-10-198-846-2230	Sequence 2230, Ap
	346	14	7.7	841	9	US-10-198-846-6291	Sequence 6291, Ap
	347	14	7.7	843	9	US-10-198-846-3630	Sequence 3630, Ap
	348	14	7.7	843	9	US-10-198-846-9227	Sequence 9227, Ap
	349	14	7.7	846	9	US-10-198-846-6315	Sequence 6315, Ap
	350	14	7.7	847	9	US-10-198-846-7517	Sequence 7517, Ap
	351	14	7.7	848	9	US-10-198-846-3707	Sequence 3707, Ap
	352	14	7.7	849	9	US-10-198-846-6151	Sequence 6151, Ap
	353	14	7.7	850	9	US-10-198-846-475	Sequence 475, App
	354	14	7.7	852	9	US-10-198-846-3703	Sequence 3703, Ap
	355	14	7.7	852	9	US-10-198-846-5034	Sequence 5034, Ap
	356	14	7.7	855	9	US-10-198-846-1325	Sequence 1325, Ap
	357	14	7.7	856	9	US-10-198-846-3702	Sequence 3702, Ap
	358	14	7.7	858	9	US-10-198-846-3321	Sequence 3321, Ap
	359	14	7.7	858	9	US-10-198-846-5020	Sequence 5020, Ap
	360	14	7.7	858	9	US-10-198-846-5097	Sequence 5097, Ap
	361	14	7.7	863	9	US-10-198-846-3520	Sequence 3520, Ap
	362	14	7.7	865	9	US-10-198-846-1731	Sequence 1731, Ap
	363	14	7.7	865	9	US-10-198-846-3063	Sequence 3063, Ap
	364	14	7.7	866	9	US-10-198-846-4336	Sequence 4336, Ap
	365	14	7.7	870	9	US-10-198-846-4336	Sequence 4336, Ap
	366	14	7.7	871	9	US-10-198-846-855	Sequence 855, App
	367	14	7.7	872	9	US-10-198-846-848	Sequence 848, App
	368	14	7.7	872	9	US-10-198-846-1173	Sequence 1173, Ap
	369	14	7.7	873	9	US-10-198-846-4897	Sequence 4897, Ap
	370	14	7.7	874	9	US-10-198-846-5021	Sequence 5021, Ap
	371	14	7.7	874	9	US-10-198-846-7203	Sequence 7203, Ap
	372	14	7.7	880	9	US-10-198-846-962	Sequence 962, App
	373	14	7.7	882	9	US-10-198-846-1238	Sequence 1238, Ap
	374	14	7.7	884	9	US-10-198-846-3668	Sequence 3668, Ap
	375	14	7.7	885	9	US-10-198-846-9562	Sequence 9562, Ap
	376	14	7.7	885	9	US-10-198-846-9463	Sequence 9463, Ap
	377	14	7.7	887	9	US-10-198-846-988	Sequence 988, App
	378	14	7.7	890	9	US-10-198-846-3770	Sequence 3770, Ap
	379	14	7.7	890	9	US-10-198-846-5551	Sequence 5551, Ap
	380	14	7.7	891	9	US-10-198-846-6076	Sequence 6076, Ap
	381	14	7.7	891	10	US-10-198-846-6348	Sequence 6348, Ap
	382	14	7.7	892	9	US-09-974-300-1372	Sequence 1372, Ap
	383	14	7.7			US-10-198-846-1342	Sequence 1342, Ap

Flag	No.			Ref	Code	Accession	Sequence
c	385	7.7	14	893	9	US-10-198-846-2479	Sequence 2479, Ap
c	386	7.7	14	894	9	US-10-198-846-2174	Sequence 2174, Ap
c	387	7.7	14	896	9	US-10-198-846-8989	Sequence 8989, Ap
c	388	7.7	14	897	9	US-10-198-846-4454	Sequence 4454, Ap
c	389	7.7	14	899	9	US-10-198-846-6083	Sequence 6083, Ap
c	390	7.7	14	900	9	US-10-198-846-3108	Sequence 3108, Ap
c	391	7.7	14	901	9	US-10-198-846-6077	Sequence 6077, Ap
c	392	7.7	14	902	9	US-10-198-846-3673	Sequence 3673, Ap
c	393	7.7	14	904	9	US-10-198-846-8332	Sequence 8332, Ap
c	394	7.7	14	905	9	US-10-198-846-8314	Sequence 8314, Ap
c	395	7.7	14	906	9	US-10-198-846-5052	Sequence 5052, Ap
c	396	7.7	14	909	9	US-10-198-846-9045	Sequence 9045, Ap
c	397	7.7	14	912	9	US-10-198-846-2887	Sequence 2887, Ap
c	398	7.7	14	912	9	US-10-198-846-5534	Sequence 5534, Ap
c	399	7.7	14	914	9	US-10-198-846-7115	Sequence 7115, Ap
c	400	7.7	14	918	9	US-10-198-846-56	Sequence 56, Appl
	401	7.7	14	918	10	US-09-815-242-4840	Sequence 4840, Ap
c	402	7.7	14	919	9	US-10-198-846-4324	Sequence 4324, Ap
c	403	7.7	14	919	9	US-10-198-846-6080	Sequence 6080, Ap
c	404	7.7	14	920	9	US-10-198-846-3385	Sequence 3385, Ap
c	405	7.7	14	924	9	US-10-198-846-57	Sequence 57, Appl
c	406	7.7	14	924	9	US-10-198-846-170	Sequence 170, App
c	407	7.7	14	924	10	US-09-917-800A-1572	Sequence 1572, Ap
c	408	7.7	14	925	9	US-10-198-846-426	Sequence 426, App
c	409	7.7	14	927	9	US-10-198-846-285	Sequence 285, App
c	410	7.7	14	927	9	US-10-198-846-4988	Sequence 4988, Ap
c	411	7.7	14	931	9	US-10-198-846-7336	Sequence 7336, Ap
c	412	7.7	14	939	9	US-10-198-846-1838	Sequence 1838, Ap
c	413	7.7	14	940	9	US-10-198-846-3681	Sequence 3681, Ap
c	414	7.7	14	943	9	US-10-198-846-10556	Sequence 10556, A
c	415	7.7	14	960	9	US-10-198-846-5845	Sequence 5845, A
c	416	7.7	14	975	9	US-10-198-846-3423	Sequence 3423, Ap
c	417	7.7	14	978	10	US-09-815-242-7446	Sequence 7446, App
c	418	7.7	14	1001	9	US-10-152-661-218	Sequence 218, App
c	419	7.7	14	1001	9	US-09-866-050A-218	Sequence 218, App
c	420	7.7	14	1015	9	US-09-866-050A-30	Sequence 30, Appl
c	421	7.7	14	1015	9	US-10-152-661-30	Sequence 30, Appl
c	422	7.7	14	1021	9	US-10-202-193-103	Sequence 103, App
c	423	7.7	14	1024	9	US-10-202-193-46	Sequence 46, Appl
c	424	7.7	14	1041	9	US-10-114-170-189	Sequence 189, App
c	425	7.7	14	1059	9	US-09-738-626-482	Sequence 482, App
c	426	7.7	14	1081	9	US-10-066-543-143	Sequence 143, App
c	427	7.7	14	1088	9	US-10-066-543-154	Sequence 154, App
	428	7.7	14	1092	10	US-09-925-101-317	Sequence 317, App
c	429	7.7	14	1113	10	US-09-738-626-481	Sequence 481, App
	430	7.7	14	1113	10	US-09-908-855-40	Sequence 40, Appl
	431	7.7	14	1116	10	US-09-974-300-2329	Sequence 2329, Ap
	432	7.7	14	1144	10	US-09-925-301-242	Sequence 242, App
	433	7.7	14	1167	10	US-09-770-445-59	Sequence 59, Appl
	434	7.7	14	1169	10	US-09-822-849A-201	Sequence 201, App
	435	7.7	14	1190	10	US-09-938-842A-1532	Sequence 1532, App
	436	7.7	14	1190	10	US-09-887-576-776	Sequence 776, Ap
c	437	7.7	14	1194	10	US-09-738-626-3016	Sequence 3016, App
c	438	7.7	14	1194	10	US-09-887-576-818	Sequence 818, App
c	439	7.7	14	1263	10	US-09-815-242-4274	Sequence 4274, Ap
	440	7.7	14	1266	10	US-09-815-242-8090	Sequence 8090, Ap
c	441	7.7	14	1266	10	US-09-815-242-8747	Sequence 8747, Ap
	442	7.7	14	1275	10	US-09-415-540-2	Sequence 2, Appli
c	443	7.7	14	1293	10	US-09-938-842A-769	Sequence 769, App
	444	7.7	14	1302	10	US-09-938-842A-2592	Sequence 2592, Ap
c	445	7.7	14	1303	9	US-09-954-531-966	Sequence 966, App
c	446	7.7	14	1303	9	US-09-954-531-1364	Sequence 1364, App
c	447	7.7	14	1312	10	US-09-765-272-69	Sequence 69, Appl
c	448	7.7	14	1350	9	US-09-984-245-115	Sequence 115, App
c	449	7.7	14	1350	9	US-09-966-262-115	Sequence 115, App
c	450	7.7	14	1350	9	US-09-983-966-115	Sequence 115, App
c	451	7.7	14	1350	9	US-10-143-090-115	Sequence 115, App
c	452	7.7	14	1351	9	US-09-769-787-322	Sequence 322, App
c	453	7.7	14	1351	9	US-09-984-245-85	Sequence 85, Appl
c	454	7.7	14	1351	9	US-09-966-262-85	Sequence 85, Appl
c	455	7.7	14	1351	9	US-09-983-966-85	Sequence 85, Appl
c	456	7.7	14	1351	9	US-10-143-090-85	Sequence 85, Appl
c	457	7.7	14	1352	10	US-09-974-300-1685	Sequence 1685, Ap

Flag	No.			Ref	Code	Accession	Sequence
c	458	14	7.7	1386	9	US-09-738-626-3023	Sequence 3023, Ap
	459	14	7.7	1405	9	US-10-114-893-199	Sequence 199, App
c	460	14	7.7	1432	9	US-10-198-846-9956	Sequence 9956, Ap
c	461	14	7.7	1442	9	US-10-144-929-63	Sequence 63, Appl
c	462	14	7.7	1497	10	US-09-938-842A-396	Sequence 396, App
	463	14	7.7	1584	10	US-09-962-436-333	Sequence 333, Appli
	464	14	7.7	1588	10	US-09-817-464-7	Sequence 7, Appli
	465	14	7.7	1594	10	US-09-915-582-24	Sequence 24, Appl
	466	14	7.7	1614	9	US-09-769-787-277	Sequence 277, App
	467	14	7.7	1614	10	US-09-815-242-9365	Sequence 9365, Ap
	468	14	7.7	1645	10	US-09-864-761-33076	Sequence 3076, A
	469	14	7.7	1673	10	US-09-915-582-44	Sequence 44, Appl
	470	14	7.7	1684	10	US-09-925-301-139	Sequence 139, App
c	471	14	7.7	1693	10	US-09-070-927A-501	Sequence 501, App
c	472	14	7.7	1728	10	US-09-938-842A-1383	Sequence 1383, Ap
c	473	14	7.7	1743	10	US-09-925-302-239	Sequence 239, App
c	474	14	7.7	1745	7	US-09-925-301-541	Sequence 541, App
	475	14	7.7	1808	10	US-08-781-986A-502	Sequence 502, App
	476	14	7.7	1826	10	US-09-925-297-192	Sequence 192, App
	477	14	7.7	1854	10	US-09-730-525-15	Sequence 15, Appl
	478	14	7.7	1854	9	US-09-730-917-15	Sequence 15, Appl
	479	14	7.7	1854	9	US-10-198-846-10908	Sequence 10908, A
	480	14	7.7	1882	10	US-10-103-313-280	Sequence 280, App
	481	14	7.7	1885	9	US-09-769-787-203	Sequence 203, App
	482	14	7.7	1887	10	US-09-954-456-1173	Sequence 1173, Ap
	483	14	7.7	1911	10	US-09-841-786-12	Sequence 12, Appl
	484	14	7.7	1915	10	US-09-917-800A-1611	Sequence 1611, Appl
	485	14	7.7	1919	9	US-09-892-985-11	Sequence 11, Appl
	486	14	7.7	1919	10	US-10-071-766-134	Sequence 134, App
	487	14	7.7	1941	10	US-09-841-132-316	Sequence 316, App
	488	14	7.7	1954	9	US-09-925-301-82	Sequence 82, Appl
c	489	14	7.7	2000	9	US-09-938-842A-2831	Sequence 2831, Ap
c	490	14	7.7	2000	9	US-09-938-842A-4030	Sequence 4030, Ap
c	491	14	7.7	2000	9	US-09-938-842A-4686	Sequence 4686, Ap
c	492	14	7.7	2000	9	US-09-938-842A-4749	Sequence 4749, Ap
c	493	14	7.7	2000	9	US-09-938-842A-5202	Sequence 5202, Ap
c	494	14	7.7	2015	9	US-09-938-842A-5311	Sequence 5311, Ap
c	495	14	7.7	2015	10	US-09-954-456-1983	Sequence 1983, App
c	496	14	7.7	2029	9	US-09-764-864-331	Sequence 331, App
c	497	14	7.7	2029	10	US-10-077-583-1	Sequence 1, Appli
c	498	14	7.7	2083	10	US-09-822-846-539	Sequence 539, App
	499	14	7.7	2086	10	US-09-908-855-37	Sequence 37, Appl
c	500	14	7.7	2109	7	US-08-781-986A-416	Sequence 416, App
c	501	14	7.7	2110	10	US-09-938-803-36	Sequence 36, Appl
c	502	14	7.7	2127	10	US-09-974-300-1668	Sequence 1668, Ap
c	503	14	7.7	2201	10	US-09-764-864-233	Sequence 233, App
c	504	14	7.7	2268	10	US-09-815-242-4341	Sequence 4341, Ap
c	505	14	7.7	2274	10	US-09-815-242-8547	Sequence 8547, Ap
c	506	14	7.7	2376	10	US-09-925-301-497	Sequence 497, App
c	507	14	7.7	2403	10	US-09-815-242-9287	Sequence 9287, App
	508	14	7.7	2468	9	US-09-764-868-1450	Sequence 1450, App
	509	14	7.7	2517	9	US-09-894-030-1	Sequence 1, Appli
c	510	14	7.7	2517	10	US-09-893-600-1	Sequence 1, Appli
c	511	14	7.7	2520	10	US-09-893-499-1	Sequence 1, Appli
c	512	14	7.7	2533	10	US-09-938-842A-1600	Sequence 1600, Ap
c	513	14	7.7	2533	10	US-09-984-271-59	Sequence 59, Appl
	514	14	7.7	2589	10	US-09-815-242-5988	Sequence 5988, Ap
c	515	14	7.7	2589	10	US-09-938-842A-1345	Sequence 1345, Ap
	516	14	7.7	2652	10	US-09-741-669-250	Sequence 250, App
	517	14	7.7	2709	9	US-09-916-790-6	Sequence 6, Appli
	518	14	7.7	2742	10	US-09-938-842A-297	Sequence 297, App
c	519	14	7.7	2816	10	US-09-815-242-4163	Sequence 4163, Ap
c	520	14	7.7	2816	10	US-10-198-846-9693	Sequence 9693, Ap
	521	14	7.7	3104	10	US-10-198-846-9882	Sequence 9882, Ap
	522	14	7.7	3107	10	US-10-037-598-32	Sequence 32, Appl
	523	14	7.7	3138	10	US-09-783-066-5	Sequence 5, Appli
	524	14	7.7	3283	9	US-09-950-041-11	Sequence 11, Appli
c	525	14	7.7	3283	9	US-09-978-295A-495	Sequence 495, App
c	526	14	7.7	3283	9	US-09-978-697-495	Sequence 495, App
c	527	14	7.7	3283	9	US-09-978-192A-495	Sequence 495, App
c	528	14	7.7	3283	9	US-09-978-832A-495	Sequence 495, App
c	529	14	7.7	3283	9	US-09-999-189-495	Sequence 495, App
c	530	14	7.7	3283	9	US-10-028-072-357	Sequence 357, App

#					Identifier	Sequence
531	14	7.7	3283	9	US-10-121-049-357	Sequence 357, App
532	14	7.7	3283	9	US-10-123-904-357	Sequence 357, App
533	14	7.7	3283	9	US-10-140-470-357	Sequence 357, App
534	14	7.7	3283	9	US-10-095-627-2	Sequence 2, Appli
535	14	7.7	3283	9	US-10-175-746-357	Sequence 357, App
536	14	7.7	3283	9	US-10-176-918-357	Sequence 357, App
537	14	7.7	3283	9	US-10-176-921-357	Sequence 357, App
538	14	7.7	3283	9	US-10-137-865-357	Sequence 357, App
539	14	7.7	3283	9	US-10-140-474-357	Sequence 357, App
540	14	7.7	3283	9	US-10-142-431-357	Sequence 357, App
541	14	7.7	3283	9	US-10-143-114-357	Sequence 357, App
542	14	7.7	3283	9	US-10-140-002-357	Sequence 357, App
543	14	7.7	3283	9	US-10-140-608A-495	Sequence 495, App
544	14	7.7	3283	9	US-10-142-419-357	Sequence 357, App
545	14	7.7	3283	9	US-09-978-191A-495	Sequence 495, App
546	14	7.7	3283	9	US-09-978-403A-495	Sequence 495, App
547	14	7.7	3283	9	US-09-978-564A-495	Sequence 495, App
548	14	7.7	3283	9	US-09-978-585A-495	Sequence 495, App
549	14	7.7	3283	9	US-10-017-081A-495	Sequence 495, App
550	14	7.7	3283	9	US-10-123-262-357	Sequence 357, App
551	14	7.7	3283	9	US-10-142-423-357	Sequence 357, App
552	14	7.7	3283	9	US-09-978-824-495	Sequence 495, App
553	14	7.7	3283	9	US-09-981-915A-495	Sequence 495, App
554	14	7.7	3283	9	US-09-999-833A-495	Sequence 495, App
555	14	7.7	3283	9	US-10-121-050-357	Sequence 357, App
556	14	7.7	3283	9	US-10-141-755-357	Sequence 357, App
557	14	7.7	3283	9	US-10-167-749-495	Sequence 495, App
558	14	7.7	3283	9	US-09-918-585A-495	Sequence 495, App
559	14	7.7	3283	9	US-10-143-032-357	Sequence 357, App
560	14	7.7	3283	9	US-09-978-423A-495	Sequence 495, App
561	14	7.7	3283	9	US-10-013-921A-495	Sequence 495, App
562	14	7.7	3283	9	US-10-123-108-357	Sequence 357, App
563	14	7.7	3283	9	US-10-123-236-357	Sequence 357, App
564	14	7.7	3283	9	US-10-123-261-357	Sequence 357, App
565	14	7.7	3283	9	US-10-140-921-357	Sequence 357, App
566	14	7.7	3283	9	US-10-140-928-357	Sequence 357, App
567	14	7.7	3283	9	US-09-978-193A-495	Sequence 495, App
568	14	7.7	3283	9	US-09-999-830A-495	Sequence 495, App
569	14	7.7	3283	9	US-10-016-177A-495	Sequence 495, App
570	14	7.7	3283	9	US-10-121-045-357	Sequence 357, App
571	14	7.7	3283	9	US-10-121-047-357	Sequence 357, App
572	14	7.7	3283	9	US-10-123-292-357	Sequence 357, App
573	14	7.7	3283	9	US-10-123-903-357	Sequence 357, App
574	14	7.7	3283	9	US-10-124-819-357	Sequence 357, App
575	14	7.7	3283	9	US-10-124-822-357	Sequence 357, App
576	14	7.7	3283	9	US-10-140-925-357	Sequence 357, App
577	14	7.7	3283	9	US-10-160-498-357	Sequence 357, App
578	14	7.7	3283	9	US-10-121-041-357	Sequence 357, App
579	14	7.7	3283	9	US-10-121-043-357	Sequence 357, App
580	14	7.7	3283	9	US-10-121-047-357	Sequence 357, App
581	14	7.7	3283	9	US-10-123-215-357	Sequence 357, App
582	14	7.7	3283	9	US-10-123-902-357	Sequence 357, App
583	14	7.7	3283	9	US-10-123-908-357	Sequence 357, App
584	14	7.7	3283	9	US-10-123-909-357	Sequence 357, App
585	14	7.7	3283	9	US-10-123-910-357	Sequence 357, App
586	14	7.7	3283	9	US-10-124-813-357	Sequence 357, App
587	14	7.7	3283	9	US-10-124-817-357	Sequence 357, App
588	14	7.7	3283	9	US-10-124-824-357	Sequence 357, App
589	14	7.7	3283	9	US-10-125-922-357	Sequence 357, App
590	14	7.7	3283	9	US-10-125-924-357	Sequence 357, App
591	14	7.7	3283	9	US-10-127-825A-357	Sequence 357, App
592	14	7.7	3283	9	US-10-127-829A-357	Sequence 357, App
593	14	7.7	3283	9	US-10-127-835A-357	Sequence 357, App
594	14	7.7	3283	9	US-10-127-839A-357	Sequence 357, App
595	14	7.7	3283	9	US-10-127-901A-357	Sequence 357, App
596	14	7.7	3283	9	US-10-128-693A-357	Sequence 357, App
597	14	7.7	3283	9	US-10-131-813A-357	Sequence 357, App
598	14	7.7	3283	9	US-10-131-818A-357	Sequence 357, App
599	14	7.7	3283	9	US-10-131-823A-357	Sequence 357, App
600	14	7.7	3283	9	US-10-131-824A-357	Sequence 357, App
601	14	7.7	3283	9	US-10-131-830A-357	Sequence 357, App
602	14	7.7	3283	9	US-10-131-837A-357	Sequence 357, App
603	14	7.7	3283	9	US-10-137-872A-357	Sequence 357, App
604	14	7.7	3283	9	US-10-140-860-357	Sequence 357, App
605	14	7.7	3283	9	US-10-142-417-357	Sequence 357, App
606	14	7.7	3283	9	US-10-147-500-357	Sequence 357, App
607	14	7.7	3283	9	US-10-147-502-357	Sequence 357, App
608	14	7.7	3283	9	US-10-147-515-357	Sequence 357, App
609	14	7.7	3283	9	US-10-147-517-357	Sequence 357, App
610	14	7.7	3283	9	US-10-147-519-357	Sequence 357, App
611	14	7.7	3283	9	US-10-147-526-357	Sequence 357, App
612	14	7.7	3283	9	US-10-147-527-357	Sequence 357, App
613	14	7.7	3283	9	US-10-152-395-357	Sequence 357, App
614	14	7.7	3283	9	US-10-157-782-357	Sequence 357, App
615	14	7.7	3283	9	US-09-978-757A-495	Sequence 495, App
616	14	7.7	3283	9	US-10-121-040-357	Sequence 357, App
617	14	7.7	3283	9	US-10-121-056-357	Sequence 357, App
618	14	7.7	3283	9	US-10-121-061-357	Sequence 357, App
619	14	7.7	3283	9	US-10-123-235-357	Sequence 357, App
620	14	7.7	3283	9	US-10-124-818-357	Sequence 357, App
621	14	7.7	3283	9	US-10-125-926A-357	Sequence 357, App
622	14	7.7	3283	9	US-10-125-930A-357	Sequence 357, App
623	14	7.7	3283	9	US-10-127-831A-357	Sequence 357, App
624	14	7.7	3283	9	US-10-127-837A-357	Sequence 357, App
625	14	7.7	3283	9	US-10-127-838B-357	Sequence 357, App
626	14	7.7	3283	9	US-10-127-842A-357	Sequence 357, App
627	14	7.7	3283	9	US-10-127-843A-357	Sequence 357, App
628	14	7.7	3283	9	US-10-127-845A-357	Sequence 357, App
629	14	7.7	3283	9	US-10-127-846A-357	Sequence 357, App
630	14	7.7	3283	9	US-10-127-848A-357	Sequence 357, App
631	14	7.7	3283	9	US-10-127-849A-357	Sequence 357, App
632	14	7.7	3283	9	US-10-127-850A-357	Sequence 357, App
633	14	7.7	3283	9	US-10-127-851A-357	Sequence 357, App
634	14	7.7	3283	9	US-10-128-684A-357	Sequence 357, App
635	14	7.7	3283	9	US-10-128-686A-357	Sequence 357, App
636	14	7.7	3283	9	US-10-128-690A-357	Sequence 357, App
637	14	7.7	3283	9	US-10-128-691A-357	Sequence 357, App
638	14	7.7	3283	9	US-10-131-819A-357	Sequence 357, App
639	14	7.7	3283	9	US-10-131-829A-357	Sequence 357, App
640	14	7.7	3283	9	US-10-131-836A-357	Sequence 357, App
641	14	7.7	3283	9	US-10-137-868-357	Sequence 357, App
642	14	7.7	3283	9	US-10-146-729-357	Sequence 357, App
643	14	7.7	3283	9	US-10-146-791-357	Sequence 357, App
644	14	7.7	3283	9	US-10-147-484-357	Sequence 357, App
645	14	7.7	3283	9	US-10-147-492-357	Sequence 357, App
646	14	7.7	3283	9	US-10-147-508-357	Sequence 357, App
647	14	7.7	3283	9	US-10-147-512-357	Sequence 357, App
648	14	7.7	3283	9	US-10-158-782-357	Sequence 357, App
649	14	7.7	3283	9	US-10-175-735-357	Sequence 357, App
650	14	7.7	3283	9	US-10-123-905-357	Sequence 357, App
651	14	7.7	3283	9	US-10-123-907-357	Sequence 357, App
652	14	7.7	3283	9	US-10-124-815-357	Sequence 357, App
653	14	7.7	3283	9	US-10-125-921A-357	Sequence 357, App
654	14	7.7	3283	9	US-10-125-928A-357	Sequence 357, App
655	14	7.7	3283	9	US-10-127-821A-357	Sequence 357, App
656	14	7.7	3283	9	US-10-127-824A-357	Sequence 357, App
657	14	7.7	3283	9	US-10-127-826A-357	Sequence 357, App
658	14	7.7	3283	9	US-10-127-827A-357	Sequence 357, App
659	14	7.7	3283	9	US-10-127-828A-357	Sequence 357, App
660	14	7.7	3283	9	US-10-127-830A-357	Sequence 357, App
661	14	7.7	3283	9	US-10-127-832A-357	Sequence 357, App
662	14	7.7	3283	9	US-10-127-833A-357	Sequence 357, App
663	14	7.7	3283	9	US-10-127-834A-357	Sequence 357, App
664	14	7.7	3283	9	US-10-127-836A-357	Sequence 357, App
665	14	7.7	3283	9	US-10-127-841A-357	Sequence 357, App
666	14	7.7	3283	9	US-10-127-844A-357	Sequence 357, App
667	14	7.7	3283	9	US-10-128-687A-357	Sequence 357, App
668	14	7.7	3283	9	US-10-128-688A-357	Sequence 357, App
669	14	7.7	3283	9	US-10-128-689A-357	Sequence 357, App
670	14	7.7	3283	9	US-10-128-694A-357	Sequence 357, App
671	14	7.7	3283	9	US-10-131-825A-357	Sequence 357, App
672	14	7.7	3283	9	US-10-230-417-357	Sequence 357, App
673	14	7.7	3283	9	US-10-235-767-2	Sequence 2, Appli
674	14	7.7	3283	9	US-10-121-051-357	Sequence 357, App
675	14	7.7	3283	9	US-10-131-815A-357	Sequence 357, App
676	14	7.7	3283	9	US-10-131-815A-357	Sequence 357, App

No.		len	score	sz	exp	Sequence ID	Label
677		14	7.7	3283	9	US-10-131-817A-357	Sequence 357, App
678		14	7.7	3283	9	US-10-131-821A-357	Sequence 357, App
679		14	7.7	3283	9	US-10-131-822A-357	Sequence 357, App
680		14	7.7	3283	9	US-10-131-828A-357	Sequence 357, App
681		14	7.7	3283	9	US-10-131-835A-357	Sequence 357, App
682		14	7.7	3283	9	US-10-137-864A-357	Sequence 357, App
683		14	7.7	3283	9	US-10-137-869A-357	Sequence 357, App
684		14	7.7	3283	9	US-10-147-523-357	Sequence 357, App
685		14	7.7	3283	9	US-10-158-785-357	Sequence 357, App
686		14	7.7	3283	9	US-09-978-187B-495	Sequence 495, App
687		14	7.7	3283	9	US-10-121-042-357	Sequence 357, App
688		14	7.7	3283	9	US-10-123-912-357	Sequence 357, App
689		14	7.7	3313	10	US-09-974-300-174	Sequence 174, App
690		14	7.7	3462	10	US-09-742-373-10	Sequence 10, Appl
691		14	7.7	3466	10	US-10-198-846-13852	Sequence 13852, A
692		14	7.7	3697	12	US-10-002-600-58	Sequence 58, Appl
693	c	14	7.7	3871	9	US-10-174-590-347	Sequence 347, App
694	c	14	7.7	3871	9	US-10-176-758-347	Sequence 347, App
695	c	14	7.7	3871	9	US-10-175-737-347	Sequence 347, App
696	c	14	7.7	3871	9	US-10-173-706-347	Sequence 347, App
697	c	14	7.7	3871	9	US-10-175-738-347	Sequence 347, App
698	c	14	7.7	3871	9	US-10-175-752-347	Sequence 347, App
699	c	14	7.7	3871	9	US-10-176-482-347	Sequence 347, App
700	c	14	7.7	3871	9	US-10-176-757-347	Sequence 347, App
701	c	14	7.7	3871	9	US-10-176-913-347	Sequence 347, App
702	c	14	7.7	3871	9	US-10-180-552-347	Sequence 347, App
703	c	14	7.7	3871	9	US-10-180-557-347	Sequence 347, App
704	c	14	7.7	3871	9	US-10-174-572-347	Sequence 347, App
705	c	14	7.7	3871	9	US-10-174-579-347	Sequence 347, App
706	c	14	7.7	3871	9	US-10-174-582-347	Sequence 347, App
707	c	14	7.7	3871	9	US-10-174-588-347	Sequence 347, App
708	c	14	7.7	3871	9	US-10-174-739-347	Sequence 347, App
709	c	14	7.7	3871	9	US-10-175-739-347	Sequence 347, App
710	c	14	7.7	3871	9	US-10-175-740-347	Sequence 347, App
711	c	14	7.7	3871	9	US-10-175-743-347	Sequence 347, App
712	c	14	7.7	3871	9	US-10-176-488-347	Sequence 347, App
713	c	14	7.7	3871	9	US-10-176-492-347	Sequence 347, App
714	c	14	7.7	3871	9	US-10-176-747-347	Sequence 347, App
715	c	14	7.7	3871	9	US-10-176-750-347	Sequence 347, App
716	c	14	7.7	3871	9	US-10-176-985-347	Sequence 347, App
717	c	14	7.7	3871	9	US-10-176-987-347	Sequence 347, App
718	c	14	7.7	3871	9	US-10-176-991-347	Sequence 347, App
719	c	14	7.7	3871	9	US-10-176-992-347	Sequence 347, App
720	c	14	7.7	3871	9	US-10-176-993-347	Sequence 347, App
721	c	14	7.7	3871	9	US-10-184-658-347	Sequence 347, App
722	c	14	7.7	3871	9	US-10-173-695-347	Sequence 347, App
723	c	14	7.7	3871	9	US-10-173-697-347	Sequence 347, App
724	c	14	7.7	3871	9	US-10-173-705-347	Sequence 347, App
725	c	14	7.7	3871	9	US-10-174-576-347	Sequence 347, App
726	c	14	7.7	3871	9	US-10-174-585-347	Sequence 347, App
727	c	14	7.7	3871	9	US-10-174-586-347	Sequence 347, App
728	c	14	7.7	3871	9	US-10-175-747-347	Sequence 347, App
729	c	14	7.7	3871	9	US-10-176-481-347	Sequence 347, App
730	c	14	7.7	3871	9	US-10-176-485-347	Sequence 347, App
731	c	14	7.7	3871	9	US-10-176-487-347	Sequence 347, App
732	c	14	7.7	3871	9	US-10-176-493-347	Sequence 347, App
733	c	14	7.7	3871	9	US-10-176-756-347	Sequence 347, App
734	c	14	7.7	3871	9	US-10-176-911-347	Sequence 347, App
735	c	14	7.7	3871	9	US-10-176-919-347	Sequence 347, App
736	c	14	7.7	3871	9	US-10-176-925-347	Sequence 347, App
737	c	14	7.7	3871	9	US-10-176-978-347	Sequence 347, App
738	c	14	7.7	3871	9	US-10-179-510-347	Sequence 347, App
739	c	14	7.7	3871	9	US-10-180-543-347	Sequence 347, App
740	c	14	7.7	3871	9	US-10-180-544-347	Sequence 347, App
741	c	14	7.7	3871	9	US-10-180-546-347	Sequence 347, App
742	c	14	7.7	3871	9	US-10-180-547-347	Sequence 347, App
743	c	14	7.7	3871	9	US-10-180-549-347	Sequence 347, App
744	c	14	7.7	3871	9	US-10-180-555-347	Sequence 347, App
745	c	14	7.7	3871	9	US-10-180-559-347	Sequence 347, App
746	c	14	7.7	3871	9	US-10-181-000-347	Sequence 347, App
747	c	14	7.7	3871	9	US-10-183-010-347	Sequence 347, App
748	c	14	7.7	3871	9	US-10-183-012-347	Sequence 347, App
749	c	14	7.7	3871	9	US-10-184-614-347	Sequence 347, App

No.		len	score	sz	exp	Sequence ID	Label
750	c	14	7.7	3871	9	US-10-184-623-347	Sequence 347, App
751	c	14	7.7	3871	9	US-10-184-635-347	Sequence 347, App
752	c	14	7.7	3871	9	US-10-184-637-347	Sequence 347, App
753	c	14	7.7	3871	9	US-10-184-646-347	Sequence 347, App
754	c	14	7.7	3871	9	US-10-184-647-347	Sequence 347, App
755	c	14	7.7	3871	9	US-10-184-652-347	Sequence 347, App
756	c	14	7.7	3871	9	US-10-187-594-347	Sequence 347, App
757	c	14	7.7	3871	9	US-10-187-596-347	Sequence 347, App
758	c	14	7.7	3871	9	US-10-187-745-347	Sequence 347, App
759	c	14	7.7	3871	9	US-10-187-885-347	Sequence 347, App
760	c	14	7.7	3871	9	US-10-187-886-347	Sequence 347, App
761	c	14	7.7	3871	9	US-10-199-464-347	Sequence 347, App
762	c	14	7.7	3871	9	US-10-176-751-347	Sequence 347, App
763	c	14	7.7	3871	9	US-10-176-760-347	Sequence 347, App
764	c	14	7.7	3871	9	US-10-180-541-347	Sequence 347, App
765	c	14	7.7	3871	9	US-10-180-542-347	Sequence 347, App
766	c	14	7.7	3871	9	US-10-180-548-347	Sequence 347, App
767	c	14	7.7	3871	9	US-10-180-551-347	Sequence 347, App
768	c	14	7.7	3871	9	US-10-180-998-347	Sequence 347, App
769	c	14	7.7	3871	9	US-10-180-999-347	Sequence 347, App
770	c	14	7.7	3871	9	US-10-183-013-347	Sequence 347, App
771	c	14	7.7	3871	9	US-10-184-612-347	Sequence 347, App
772	c	14	7.7	3871	9	US-10-184-616-347	Sequence 347, App
773	c	14	7.7	3871	9	US-10-184-617-347	Sequence 347, App
774	c	14	7.7	3871	9	US-10-184-622-347	Sequence 347, App
775	c	14	7.7	3871	9	US-10-184-628-347	Sequence 347, App
776	c	14	7.7	3871	9	US-10-184-629-347	Sequence 347, App
777	c	14	7.7	3871	9	US-10-184-630-347	Sequence 347, App
778	c	14	7.7	3871	9	US-10-184-631-347	Sequence 347, App
779	c	14	7.7	3871	9	US-10-184-632-347	Sequence 347, App
780	c	14	7.7	3871	9	US-10-184-636-347	Sequence 347, App
781	c	14	7.7	3871	9	US-10-184-640-347	Sequence 347, App
782	c	14	7.7	3871	9	US-10-184-650-347	Sequence 347, App
783	c	14	7.7	3871	9	US-10-184-651-347	Sequence 347, App
784	c	14	7.7	3871	9	US-10-187-588-347	Sequence 347, App
785	c	14	7.7	3871	9	US-10-187-597-347	Sequence 347, App
786	c	14	7.7	3871	9	US-10-187-598-347	Sequence 347, App
787	c	14	7.7	3871	9	US-10-187-600-347	Sequence 347, App
788	c	14	7.7	3871	9	US-10-187-601-347	Sequence 347, App
789	c	14	7.7	3871	9	US-10-187-603-347	Sequence 347, App
790	c	14	7.7	3871	9	US-10-187-741-347	Sequence 347, App
791	c	14	7.7	3871	9	US-10-187-743-347	Sequence 347, App
792	c	14	7.7	3871	9	US-10-187-746-347	Sequence 347, App
793	c	14	7.7	3871	9	US-10-187-747-347	Sequence 347, App
794	c	14	7.7	3871	9	US-10-187-751-347	Sequence 347, App
795	c	14	7.7	3871	9	US-10-187-753-347	Sequence 347, App
796	c	14	7.7	3871	9	US-10-187-754-347	Sequence 347, App
797	c	14	7.7	3871	9	US-10-187-757-347	Sequence 347, App
798	c	14	7.7	3871	9	US-10-187-884-347	Sequence 347, App
799	c	14	7.7	3871	9	US-10-188-767-347	Sequence 347, App
800	c	14	7.7	3871	9	US-10-188-769-347	Sequence 347, App
801	c	14	7.7	3871	9	US-10-188-773-347	Sequence 347, App
802	c	14	7.7	3871	9	US-10-188-781-347	Sequence 347, App
803	c	14	7.7	3871	9	US-10-194-361-347	Sequence 347, App
804	c	14	7.7	3871	9	US-10-194-423-347	Sequence 347, App
805	c	14	7.7	3871	9	US-10-195-897-347	Sequence 347, App
806	c	14	7.7	3871	9	US-10-195-901-347	Sequence 347, App
807	c	14	7.7	3871	9	US-10-196-756-347	Sequence 347, App
808	c	14	7.7	3871	9	US-10-172-708-347	Sequence 347, App
809	c	14	7.7	3871	9	US-10-176-479-347	Sequence 347, App
810	c	14	7.7	3871	9	US-10-176-748-347	Sequence 347, App
811	c	14	7.7	3871	9	US-10-176-916-347	Sequence 347, App
812	c	14	7.7	3871	9	US-10-179-507-347	Sequence 347, App
813	c	14	7.7	3871	9	US-10-179-516-347	Sequence 347, App
814	c	14	7.7	3871	9	US-10-179-519-347	Sequence 347, App
815	c	14	7.7	3871	9	US-10-179-525-347	Sequence 347, App
816	c	14	7.7	3871	9	US-10-180-540-347	Sequence 347, App
817	c	14	7.7	3871	9	US-10-180-545-347	Sequence 347, App
818	c	14	7.7	3871	9	US-10-183-006-347	Sequence 347, App
819	c	14	7.7	3871	9	US-10-183-008-347	Sequence 347, App
820	c	14	7.7	3871	9	US-10-183-008-347	Sequence 347, App
821	c	14	7.7	3871	9	US-10-183-008-347	Sequence 347, App
822	c	14	7.7	3871	9	US-10-183-008-347	Sequence 347, App

All rows share the constant columns: `14` `7.7` `3871` `9`.

Left column

No.	Identifier	Sequence
c 823	US-10-183-017-347	Sequence 347, App
c 824	US-10-183-019-347	Sequence 347, App
c 825	US-10-184-618-347	Sequence 347, App
c 826	US-10-184-625-347	Sequence 347, App
c 827	US-10-184-626-347	Sequence 347, App
c 828	US-10-184-645-347	Sequence 347, App
c 829	US-10-184-654-347	Sequence 347, App
c 830	US-10-184-655-347	Sequence 347, App
c 831	US-10-188-774-347	Sequence 347, App
c 832	US-10-188-775-347	Sequence 347, App
c 833	US-10-194-462-347	Sequence 347, App
c 834	US-10-195-902-347	Sequence 347, App
c 835	US-10-196-743-347	Sequence 347, App
c 836	US-10-196-760-347	Sequence 347, App
c 837	US-10-196-762-347	Sequence 347, App
c 838	US-10-197-695-347	Sequence 347, App
c 839	US-10-006-856A-309	Sequence 309, App
c 840	US-10-176-484-347	Sequence 347, App
c 841	US-10-176-753-347	Sequence 347, App
c 842	US-10-176-917-347	Sequence 347, App
c 843	US-10-179-506-347	Sequence 347, App
c 844	US-10-179-513-347	Sequence 347, App
c 845	US-10-179-514-347	Sequence 347, App
c 846	US-10-179-522-347	Sequence 347, App
c 847	US-10-180-556-347	Sequence 347, App
c 848	US-10-180-560-347	Sequence 347, App
c 849	US-10-183-015-347	Sequence 347, App
c 850	US-10-184-615-347	Sequence 347, App
c 851	US-10-184-620-347	Sequence 347, App
c 852	US-10-184-643-347	Sequence 347, App
c 853	US-10-184-656-347	Sequence 347, App
c 854	US-10-192-010-347	Sequence 347, App
c 855	US-10-195-894-347	Sequence 347, App
c 856	US-10-205-908-347	Sequence 347, App
c 857	US-10-186-855-347	Sequence 347, App
c 858	US-10-187-599-347	Sequence 347, App
c 859	US-10-187-750-347	Sequence 347, App
c 860	US-10-188-780-347	Sequence 347, App
c 861	US-10-192-015-347	Sequence 347, App
c 862	US-10-194-394-347	Sequence 347, App
c 863	US-10-194-425-347	Sequence 347, App
c 864	US-10-194-485-347	Sequence 347, App
c 865	US-10-195-885-347	Sequence 347, App
c 866	US-10-195-890-347	Sequence 347, App
c 867	US-10-195-899-347	Sequence 347, App
c 868	US-10-196-748-347	Sequence 347, App
c 869	US-10-196-750-347	Sequence 347, App
c 870	US-10-197-699-347	Sequence 347, App
c 871	US-10-197-700-347	Sequence 347, App
c 872	US-10-197-705-347	Sequence 347, App
c 873	US-10-197-708-347	Sequence 347, App
c 874	US-10-198-764-347	Sequence 347, App
c 875	US-10-198-765-347	Sequence 347, App
c 876	US-10-198-768-347	Sequence 347, App
c 877	US-10-198-769-347	Sequence 347, App
c 878	US-10-199-305-347	Sequence 347, App
c 879	US-10-199-306-347	Sequence 347, App
c 880	US-10-199-310-347	Sequence 347, App
c 881	US-10-199-311-347	Sequence 347, App
c 882	US-10-199-314-347	Sequence 347, App
c 883	US-10-199-317-347	Sequence 347, App
c 884	US-10-199-665-347	Sequence 347, App
c 885	US-10-199-666-347	Sequence 347, App
c 886	US-10-199-669-347	Sequence 347, App
c 887	US-10-201-534-347	Sequence 347, App
c 888	US-10-201-770-347	Sequence 347, App
c 889	US-10-201-855-347	Sequence 347, App
c 890	US-10-201-856-347	Sequence 347, App
c 891	US-10-202-469-347	Sequence 347, App
c 892	US-10-202-470-347	Sequence 347, App
c 893	US-10-202-...	Sequence 347, App
c 894	US-10-202-...	Sequence 347, App
c 895	US-10-202-...	Sequence 347, App

Right column

No.	Identifier	Sequence
c 896	US-10-202-476-347	Sequence 347, App
c 897	US-10-202-933-347	Sequence 347, App
c 898	US-10-202-935-347	Sequence 347, App
c 899	US-10-202-936-347	Sequence 347, App
c 900	US-10-202-939-347	Sequence 347, App
c 901	US-10-205-504-347	Sequence 347, App
c 902	US-10-205-509-347	Sequence 347, App
c 903	US-10-205-895-347	Sequence 347, App
c 904	US-10-205-899-347	Sequence 347, App
c 905	US-10-205-900-347	Sequence 347, App
c 906	US-10-205-909-347	Sequence 347, App
c 907	US-10-006-818A-309	Sequence 309, App
c 908	US-10-183-002-347	Sequence 347, App
c 909	US-10-184-621-347	Sequence 347, App
c 910	US-10-184-638-347	Sequence 347, App
c 911	US-10-187-752-347	Sequence 347, App
c 912	US-10-187-887-347	Sequence 347, App
c 913	US-10-194-461-347	Sequence 347, App
c 914	US-10-195-892-347	Sequence 347, App
c 915	US-10-195-751-347	Sequence 347, App
c 916	US-10-197-694-347	Sequence 347, App
c 917	US-10-197-697-347	Sequence 347, App
c 918	US-10-197-707-347	Sequence 347, App
c 919	US-10-199-303-347	Sequence 347, App
c 920	US-10-199-318-347	Sequence 347, App
c 921	US-10-199-458-347	Sequence 347, App
c 922	US-10-199-462-347	Sequence 347, App
c 923	US-10-201-324-347	Sequence 347, App
c 924	US-10-201-328-347	Sequence 347, App
c 925	US-10-201-527-347	Sequence 347, App
c 926	US-10-201-528-347	Sequence 347, App
c 927	US-10-201-529-347	Sequence 347, App
c 928	US-10-201-530-347	Sequence 347, App
c 929	US-10-202-408-347	Sequence 347, App
c 930	US-10-202-409-347	Sequence 347, App
c 931	US-10-202-411-347	Sequence 347, App
c 932	US-10-202-472-347	Sequence 347, App
c 933	US-10-205-502-347	Sequence 347, App
c 934	US-10-205-507-347	Sequence 347, App
c 935	US-10-205-511-347	Sequence 347, App
c 936	US-10-205-902-347	Sequence 347, App
c 937	US-10-205-907-347	Sequence 347, App
c 938	US-10-194-456-347	Sequence 347, App
c 939	US-10-196-758-347	Sequence 347, App
c 940	US-10-199-770-347	Sequence 347, App
c 941	US-10-199-308-347	Sequence 347, App
c 942	US-10-200-617-347	Sequence 347, App
c 943	US-10-205-893-347	Sequence 347, App
c 944	US-10-205-897-347	Sequence 347, App
c 945	US-10-015-393A-309	Sequence 309, App
c 946	US-10-174-571-347	Sequence 347, App
c 947	US-10-176-746-347	Sequence 347, App
c 948	US-10-176-923-347	Sequence 347, App
c 949	US-10-183-011-347	Sequence 347, App
c 950	US-10-184-633-347	Sequence 347, App
c 951	US-10-184-639-347	Sequence 347, App
c 952	US-10-187-742-347	Sequence 347, App
c 953	US-10-187-748-347	Sequence 347, App
c 954	US-10-188-766-347	Sequence 347, App
c 955	US-10-188-771-347	Sequence 347, App
c 956	US-10-192-006-347	Sequence 347, App
c 957	US-10-192-008-347	Sequence 347, App
c 958	US-10-192-009-347	Sequence 347, App
c 959	US-10-192-012-347	Sequence 347, App
c 960	US-10-192-016-347	Sequence 347, App
c 961	US-10-194-362-347	Sequence 347, App
c 962	US-10-194-364-347	Sequence 347, App
c 963	US-10-194-395-347	Sequence 347, App
c 964	US-10-194-424-347	Sequence 347, App
c 965	US-10-194-458-347	Sequence 347, App
c 966	US-10-194-459-347	Sequence 347, App
c 967	US-10-194-488-347	Sequence 347, App
c 968	US-10-194-488-347	Sequence 347, App

```
c 969   14  7.7  3871  9  US-10-195-886-347    Sequence 347, App
c 970   14  7.7  3871  9  US-10-195-891-347    Sequence 347, App
c 971   14  7.7  3871  9  US-10-196-746-347    Sequence 347, App
c 972   14  7.7  3871  9  US-10-196-752-347    Sequence 347, App
c 973   14  7.7  3871  9  US-10-196-753-347    Sequence 347, App
c 974   14  7.7  3871  9  US-10-196-754-347    Sequence 347, App
c 975   14  7.7  3871  9  US-10-196-761-347    Sequence 347, App
c 976   14  7.7  3871  9  US-10-197-692-347    Sequence 347, App
c 977   14  7.7  3871  9  US-10-197-693-347    Sequence 347, App
c 978   14  7.7  3871  9  US-10-197-696-347    Sequence 347, App
c 979   14  7.7  3871  9  US-10-197-698-347    Sequence 347, App
c 980   14  7.7  3871  9  US-10-197-703-347    Sequence 347, App
c 981   14  7.7  3871  9  US-10-197-711-347    Sequence 347, App
c 982   14  7.7  3871  9  US-10-198-757-347    Sequence 347, App
c 983   14  7.7  3871  9  US-10-198-761-347    Sequence 347, App
c 984   14  7.7  3871  9  US-10-198-762-347    Sequence 347, App
c 985   14  7.7  3871  9  US-10-198-763-347    Sequence 347, App
c 986   14  7.7  3871  9  US-10-198-767-347    Sequence 347, App
c 987   14  7.7  3871  9  US-10-199-301-347    Sequence 347, App
c 988   14  7.7  3871  9  US-10-199-307-347    Sequence 347, App
c 989   14  7.7  3871  9  US-10-199-312-347    Sequence 347, App
c 990   14  7.7  3871  9  US-10-199-315-347    Sequence 347, App
c 991   14  7.7  3871  9  US-10-199-316-347    Sequence 347, App
c 992   14  7.7  3871  9  US-10-199-457-347    Sequence 347, App
c 993   14  7.7  3871  9  US-10-199-459-347    Sequence 347, App
c 994   14  7.7  3871  9  US-10-199-460-347    Sequence 347, App
c 995   14  7.7  3871  9  US-10-199-461-347    Sequence 347, App
c 996   14  7.7  3871  9  US-10-199-667-347    Sequence 347, App
c 997   14  7.7  3871  9  US-10-199-673-347    Sequence 347, App
c 998   14  7.7  3871  9  US-10-201-321-347    Sequence 347, App
c 999   14  7.7  3871  9  US-10-201-322-347    Sequence 347, App
c1000   14  7.7  3871  9  US-10-201-326-347    Sequence 347, App
```

ALIGNMENTS

RESULT 1
US-09-864-761-33139
; Sequence 33139, Application US2002004876361
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33139
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000120.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
; OTHER INFORMATION: EST_HUMAN HIT: AI654552.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y6J6, EVALUE 8.00e-67
; OTHER INFORMATION: NT HIT: gi|15426220, EVALUE 0.00e+00
US-09-864-761-33139

Query Match   100.0%; Score 181; DB 10; Length 372;
Best Local Similarity   100.0%; Pred. No. 1.3e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGAGTTTTATTACTTAT  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ACTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGAGTTTTATTACTTAT  66

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 126

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATGAATGTTCTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATGAATGTTCTCTTTC 186

Qy  181 A 181
          |
Db  187 A 187
```

RESULT 2
US-09-864-761-3463
; Sequence 3463, Application US/09864761
; Patent No. US2002004876361
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359

APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16671
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000120.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
US-09-864-761-16671

Query Match   100.0%;   Score 181; DB 10;   Length 471;
Best Local Similarity   100.0%; Pred. No. 1.3e-90;
Matches 181;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ACTTTATCCAATTCACACAGACCGTCTTCCGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   245 ACTTTATCCAATTCACACAGACCGTCTTCCGAAGACGTCTTCCGAAGGATTTTATTACTTAT 304

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGCCCTCCAAGCCAAAGTTGAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   305 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGCCCTCCAAGCCAAAGTTGAT 364

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   365 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 424

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3463
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000052.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.88
; NUMBER OF SEQ ID NOS: 49117
US-09-864-761-3463

Query Match   100.0%;   Score 181; DB 10;   Length 450;
Best Local Similarity   100.0%; Pred. No. 1.3e-90;
Matches 181;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ACTTTATCCAATTCACACAGACCGTCTTCCGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 ACTTTATCCAATTCACACAGACCGTCTTCCGAAGACGTCTTCCGAAGGATTTTATTACTTAT 221

Qy   61 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGCCCTCCAAGCCAAAGTTGAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 ATGGACAATTGGCGCCAGAACACAACAGCTGAGCAAGAGCCCTCCAAGCCAAAGTTGAT 281

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 341

Qy   181 A 181
             |
Db   342 A 342

RESULT 3
US-09-864-761-16671
; Sequence 16671, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

Qy   181 A 181
Db   425 A 425

RESULT 4
US-10-000-151B-5
; Sequence 5, Application US/10000151B
; Publication No. US2003013136A1
; GENERAL INFORMATION:
; APPLICANT: Balser, Jeffrey R.
; APPLICANT: George, Alfred L.
; TITLE OF INVENTION: HUMAN KCR1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; FILE REFERENCE: Vanderbilt Ref No. US2003013136A1 VU0120; Attorney Docket No. US2003
; CURRENT APPLICATION NUMBER: US/10/000,151B
; CURRENT FILING DATE: 2000-10-30
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 5
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-000-151B-5

Query Match   100.0%; Score 181; DB 9; Length 732;
Best Local Similarity   100.0%; Pred. No. 1.3e-90;
  Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 199

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 259

Qy   181 A 181
              |
Db   260 A 260

RESULT 5
US-10-227-195A-1
; Sequence 1, Application US/10227195A
; Publication No. US2003077633A1
; GENERAL INFORMATION:
; APPLICANT: Cox, David
; APPLICANT: Arnold, Deana
; TITLE OF INVENTION: Haplotype structure of chromosome 21
; FILE REFERENCE: 1030U1
; CURRENT APPLICATION NUMBER: US/10/227,195A
; CURRENT FILING DATE: 2002-11-18
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 113604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7175_7204, 36973, 66372, 76921, 81512, 88727
; OTHER INFORMATION: n = G or C
US-10-227-195A-1

Query Match   100.0%; Score 181; DB 9; Length 113604;
Best Local Similarity   100.0%; Pred. No. 1.8e-90;
  Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60

Db   17482 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 17541

Qy   61 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17542 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 17601

Qy   121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17602 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 17661

Qy   181 A 181
              |
Db   17662 A 17662

RESULT 6
US-10-227-195A-2
; Sequence 2, Application US/10227195A
; Publication No. US2003077633A1
; GENERAL INFORMATION:
; APPLICANT: Cox, David
; APPLICANT: Arnold, Deana
; TITLE OF INVENTION: Haplotype structure of chromosome 21
; FILE REFERENCE: 1030U1
; CURRENT APPLICATION NUMBER: US/10/227,195A
; CURRENT FILING DATE: 2002-11-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 113604
; TYPE: DNA
; ORGANISM: Human
US-10-227-195A-2

Query Match   100.0%; Score 181; DB 9; Length 113604;
Best Local Similarity   100.0%; Pred. No. 1.8e-90;
  Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17482 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 17541

Qy   61 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17542 ATGGACAATTGGCGCGCCAGAACACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 17601

Qy   61 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17602 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 17661

Qy   181 A 181
              |
Db   17662 A 17662

RESULT 7
US-09-864-761-20233
; Sequence 20233, Application US/09864761
; Patent No. US2002004B763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20233
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000052.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.88
; OTHER INFORMATION: EST HUMAN HIT: AI246239.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y6J6, EVALUE 3.00e-55
; OTHER INFORMATION: NT HIT: AF302095.1, EVALUE 0.00e+00
US-09-864-761-20233

Query Match　　　74.0%;　Score 134;　DB 10;　Length 312;
Best Local Similarity　100.0%;　Pred. No. 1.7e-64;
Matches　134;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　48 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 107
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　　1 TTTTATTACTTATATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCA 60

Qy　 108 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 167
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 61 AGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGG 120

Qy　 168 AATGTTCTCTTTCA 181
　　　　　||||||||||||||
Db　　121 AATGTTCTCTTTCA 134.

RESULT 8
US-09-867-701-10563
; Sequence 10563, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10563
;   LENGTH: 714
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-10563

Query Match　　　9.9%;　Score 18;　DB 10;　Length 714;
Best Local Similarity　100.0%;　Pred. No. 5.1;
Matches　18;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　13 TTCACACAGACGCTGGAA 30
　　　　　||||||||||||||||||
Db　　84 TTCACACAGACGCTGGAA 101

RESULT 9
US-09-772-134B-89/c
; Sequence 89, Application US/09772134B
; Patent No. US2002014310A1
; GENERAL INFORMATION:
; APPLICANT: Southern Illinois University
; APPLICANT: Lightfoot, David
; APPLICANT: Meksem, Khalid
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDES AND POLYPEPTIDES RELATING TO LOCI UNDER
; TITLE OF INVENTION: RESISTANCE TO SOYBEAN CYST NEMATODE AND SOYBEAN SUDDEN DEATH SY
; TITLE OF INVENTION: METHODS EMPLOYING SAME
; FILE REFERENCE: 1268/4/2
; CURRENT APPLICATION NUMBER: US/09/772,134B
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/178,811
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89
;   LENGTH: 782
;   TYPE: DNA
;   ORGANISM: soybean
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(782)
;   OTHER INFORMATION: n is an undetermined nucleotide (dATP, dCTP, dGTP, or dTTP)
US-09-772-134B-89

Query Match　　　9.9%;　Score 18;　DB 10;　Length 782;
Best Local Similarity　100.0%;　Pred. No. 5.2;
Matches　18;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　42 AAGGATTTTATTACTTA 59
　　　　　|||||||||||||||||
Db　　781 AAGGATTTTATTACTTA 764

RESULT 10
US-09-867-701-7019
; Sequence 7019, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.

; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7019
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-7019

Query Match   9.4%;   Score 17; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 18;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   103 CTCCAAGCCAAAGTTGA 119
              |||||||||||||||||
Db   122 CTCCAAGCCAAAGTTGA 138

RESULT 11
US-09-867-701-7022
; Sequence 7022, Application US09867701
; Patent No. US2002013223A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7022
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-7022

Query Match   9.4%;   Score 17; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 18;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   103 CTCCAAGCCAAAGTTGA 119
              |||||||||||||||||
Db   122 CTCCAAGCCAAAGTTGA 138

RESULT 12
US-09-918-995-69
; Sequence 69, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(475)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-69

Query Match   9.4%;   Score 17; DB 9; Length 475;
Best Local Similarity 100.0%; Pred. No. 18;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   91 GAGCAAGAGGCCCTCCA 107
              |||||||||||||||||
Db   204 GAGCAAGAGGCCCTCCA 220

RESULT 13
US-09-918-995-1925
; Sequence 1925, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1925
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(491)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1925

Query Match   9.4%;   Score 17; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 18;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   91 GAGCAAGAGGCCCTCCA 107
              |||||||||||||||||
Db   168 GAGCAAGAGGCCCTCCA 184

RESULT 14
US-09-778-171-1/c
; Sequence 1, Application US/09778171
; Patent No. US2002012984A1
; GENERAL INFORMATION:
; APPLICANT: Mandala, et al.
; TITLE OF INVENTION: MAMMALIAN SPHINGOSINE-1-PHOSPHATE
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 20568
; CURRENT APPLICATION NUMBER: US/09/778,171
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,534
; PRIOR FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Mouse
US-09-778-171-1

Query Match   9.4%;   Score 17; DB 10; Length 2197;
Best Local Similarity 100.0%; Pred. No. 20;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     15 CACACAGAGCGCTGGAAG 31
          ||||||||||||||||||
Db    248 CACACAGAGCGCTGGAAG 232
```

RESULT 15
US-09-764-891-7358
; Sequence 7358, Application US/09764891
; Publication No. US2003007808A1
; GENERAL INFORMATION:
;   APPLICANT: Roben et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC006
;   CURRENT APPLICATION NUMBER: US/09/764,891
;   CURRENT FILING DATE: 2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 10231
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7358
;   LENGTH: 32189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7358

Query Match   9.4%; Score 17; DB 9; Length 32189;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    115 GTTGATGCTGAGAACTT 131
          |||||||||||||||||
Db   6846 GTTGATGCTGAGAACTT 6862
```

RESULT 16
US-09-854-883-243/c
; Sequence 243, Application US/09854883
; Patent No. US2002005479A1
; GENERAL INFORMATION:
;   APPLICANT: Lex M. Cowsert
;   APPLICANT: Jacqueline Wyatt
;   APPLICANT: Susan M. Freier
;   APPLICANT: Brett P. Monia
;   APPLICANT: Madeline M. Butler
;   APPLICANT: Robert McKay
;   TITLE OF INVENTION: ANTISENSE MODULATION OF PTP1B EXPRESSION
;   FILE REFERENCE: ISPH-0576
;   CURRENT APPLICATION NUMBER: US/09/854,883
;   CURRENT FILING DATE: 2001-05-14
;   PRIOR APPLICATION NUMBER: US 09/629,644
;   PRIOR FILING DATE: 2000-07-31
;   PRIOR APPLICATION NUMBER: US 09/487,368
;   PRIOR FILING DATE: 2000-01-18
;   NUMBER OF SEQ ID NOS: 389
; SEQ ID NO 243
;   LENGTH: 75899
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-854-883-243

Query Match   9.4%; Score 17; DB 10; Length 75899;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     42 AAGGATTTTATTACTT 58
          ||||||||||||||||
Db   6963 AAGGATTTTATTACTT 6947
```

RESULT 17
US-09-933-267A-1
; Sequence 1, Application US/09933267A
; Patent No. US20020123095A1
; GENERAL INFORMATION:
;   APPLICANT: Kalush, Francis et al.
;   TITLE OF INVENTION: Estrogen receptor alpha variants and
;   TITLE OF INVENTION: methods of detection thereof
;   FILE REFERENCE: CL000258CI4
;   CURRENT APPLICATION NUMBER: US/09/933,267A
;   CURRENT FILING DATE: 2001-08-21
;   PRIOR APPLICATION NUMBER: 60/160626
;   PRIOR FILING DATE: 1999-10-20
;   PRIOR APPLICATION NUMBER: 60/183756
;   PRIOR FILING DATE: 2000-02-22
;   PRIOR APPLICATION NUMBER: 09/692414
;   PRIOR FILING DATE: 2000-10-20
;   PRIOR APPLICATION NUMBER: 09/768184
;   PRIOR FILING DATE: 2001-01-24
;   PRIOR APPLICATION NUMBER: 09/804076
;   PRIOR FILING DATE: 2001-03-13
;   PRIOR APPLICATION NUMBER: 09/826314
;   PRIOR FILING DATE: 2001-04-05
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 465237
;   TYPE: DNA
;   ORGANISM: human
US-09-933-267A-1

Query Match   9.4%; Score 17; DB 10; Length 465237;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     46 ATTTTATTACTTATAT 62
          ||||||||||||||||
Db 389257 ATTTTATTACTTATAT 389273
```

RESULT 18
US-10-067-514-1/c
; Sequence 1, Application US/10067514
; Publication No. US2003005453A1
; GENERAL INFORMATION:
;   APPLICANT: Gretarsdottir, Solveig
;   APPLICANT: Jonsdottir, Sif
;   APPLICANT: Reynisdottir, Sigridur Th.
;   TITLE OF INVENTION: HUMAN STROKE GENE
;   FILE REFERENCE: 2345.2010-003
;   CURRENT APPLICATION NUMBER: US/10/067,514
;   CURRENT FILING DATE: 2002-02-04
;   PRIOR APPLICATION NUMBER: US 09/811/352
;   PRIOR FILING DATE: 2001-03-19
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1691139
;   TYPE: DNA
;   ORGANISM: Human
US-10-067-514-1

Query Match   9.4%; Score 17; DB 9; Length 1691139;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     41 GAAGGATTTTTATTACT 57
          |||||||||||||||||
Db 690407 GAAGGATTTTTATTACT 690391
```

RESULT 19
US-09-960-352-12908/c
; Sequence 12908, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:

```
;   APPLICANT: Warren, Wesley C.
;   APPLICANT: Tao, Nengbing
;   APPLICANT: Byatt, John C.
;   APPLICANT: Mathialagan, Nagappan
;   TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;   TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
;   FILE REFERENCE: 16511.006/37-21(10298)C
;   CURRENT APPLICATION NUMBER: US/09/960,352
;   CURRENT FILING DATE: 2001-09-24
;   NUMBER OF SEQ ID NOS: 15112
;   SEQ ID NO 12908
;   LENGTH: 324
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (237)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: 55-LIB3058-049-Q1-K1-F4
US-09-960-352-12908

  Query Match             8.8%;  Score 16;  DB 10;  Pred. No. 63;  Length 324;
  Best Local Similarity   100.0%;
  Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          45 GATTTTTATTACTTAT 60
               ||||||||||||||||
Db         316 GATTTTTATTACTTAT 301


RESULT 20
US-09-983-965-2960
; Sequence 2960, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;   APPLICANT: Warren, Wesley C.
;   APPLICANT: Tao, Nengbing
;   APPLICANT: Byatt, John C.
;   APPLICANT: Mathialagan, Nagappan
;   TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;   TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
;   FILE REFERENCE: 37-21(10297)C
;   CURRENT APPLICATION NUMBER: US/09/983,965
;   CURRENT FILING DATE: 2001-10-26
;   PRIOR APPLICATION NUMBER: US 09/465,231
;   PRIOR FILING DATE: 1999-12-15
;   PRIOR APPLICATION NUMBER: US 60/113,678
;   PRIOR FILING DATE: 1998-12-17
;   NUMBER OF SEQ ID NOS: 5912
;   SEQ ID NO 2960
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (325), (425)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: 20-LIB3058-009-Q1-K1-E7
US-09-983-965-2960

  Query Match             8.8%;  Score 16;  DB 10;  Pred. No. 64;  Length 438;
  Best Local Similarity   100.0%;
  Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          78 GAACACAACAGCTGAG 93
               ||||||||||||||||
Db          48 GAACACAACAGCTGAG 63


RESULT 21
US-09-879-536-381/c
; Sequence 381, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
;   APPLICANT: Endege, Wilson O.
;   APPLICANT: Steinmann, Kathleen E.
;   APPLICANT: Astle, Jon H.
;   APPLICANT: Burgess, Christopher C.
;   APPLICANT: Bushnell, Steven E.
;   APPLICANT: Carroll III, Eddie
;   APPLICANT: Catino, Theodore J.
;   APPLICANT: Derti, Adnan
;   APPLICANT: Ford, Donna M.
;   APPLICANT: Lewis, Marcia E.
;   APPLICANT: Monahan, John E.
;   APPLICANT: Schlegel, Robert
;   TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;   TITLE OF INVENTION:  PRODUCTS
;   FILE REFERENCE: CCD-257 (US)
;   CURRENT APPLICATION NUMBER: US/09/879,536
;   CURRENT FILING DATE: 2001-09-21
;   PRIOR APPLICATION NUMBER: US 60/088,801
;   PRIOR FILING DATE: 1998-06-10
;   NUMBER OF SEQ ID NOS: 850
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 381
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(622)
;   OTHER INFORMATION: n = A,T,C or G
US-09-879-536-381

  Query Match             8.8%;  Score 16;  DB 10;  Pred. No. 66;  Length 622;
  Best Local Similarity   100.0%;
  Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          57 TTATATGGACAATTGG 72
               ||||||||||||||||
Db          22 TTATATGGACAATTGG 7


RESULT 22
US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Patent No. US2002016O378A1
; GENERAL INFORMATION:
;   APPLICANT: Harper, Jeff
;   APPLICANT: Kreps, Joel
;   APPLICANT: Wang, Xun
;   APPLICANT: Zhu, Tong
;   TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;   TITLE OF INVENTION: SAME, AND METHODS OF USE
;   FILE REFERENCE: SCRIP1300-3
;   CURRENT APPLICATION NUMBER: US/09/938,842A
;   CURRENT FILING DATE: 2001-08-24
;   PRIOR APPLICATION NUMBER: US 60/227,866
;   PRIOR FILING DATE: 2000-08-24
;   PRIOR APPLICATION NUMBER: US 60/264,647
;   PRIOR FILING DATE: 2001-01-16
;   PRIOR APPLICATION NUMBER: US 60/300,111
;   PRIOR FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 5379
;   SEQ ID NO 1502
;   LENGTH: 1977
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

  Query Match             8.8%;  Score 16;  DB 9;  Pred. No. 71;  Length 1977;
  Best Local Similarity   100.0%;
  Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         113 AAGTGATGCTGAGAA 128
```

```
Db      458 AAGGTTGATGCTGAGAA 473
            |||||||||||||||||

RESULT 23
US-09-954-456-1174
; Sequence 1174, Application US/09954456
; Patent No. US2002011505741
; GENERAL INFORMATION:
;   APPLICANT: Young, Paul
;   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;   TITLE OF INVENTION: Sets
;   FILE REFERENCE: 689290-76
;   CURRENT APPLICATION NUMBER: US/09/954,456
;   CURRENT FILING DATE: 2001-09-18
;   PRIOR APPLICATION NUMBER: US/60/233,617
;   PRIOR FILING DATE: 2000-09-18
;   PRIOR APPLICATION NUMBER: US/60/234,052
;   PRIOR FILING DATE: 2000-09-20
;   PRIOR APPLICATION NUMBER: US/60/234,923
;   PRIOR FILING DATE: 2000-09-25
;   PRIOR APPLICATION NUMBER: US/60/235,637
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,638
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,711
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,720
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,840
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,863
;   PRIOR FILING DATE: 2000-09-27
;   NUMBER OF SEQ ID NOS: 2276
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1174
;   LENGTH: 2244
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-1174

Query Match             8.8%;  Score 16;  DB 10;  Length 2244;
Best Local Similarity  100.0%;  Pred. No. 71;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       112 AAAGTTGATGCTGAGA 127
             |||||||||||||||
Db       666 AAAGTTGATGCTGAGA 681

RESULT 24
US-09-954-456-1838
; Sequence 1838, Application US/09954456
; Patent No. US2002011505741
; GENERAL INFORMATION:
;   APPLICANT: Young, Paul
;   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;   TITLE OF INVENTION: Sets
;   FILE REFERENCE: 689290-76
;   CURRENT APPLICATION NUMBER: US/09/954,456
;   CURRENT FILING DATE: 2001-09-18
;   PRIOR APPLICATION NUMBER: US/60/233,617
;   PRIOR FILING DATE: 2000-09-18
;   PRIOR APPLICATION NUMBER: US/60/234,052
;   PRIOR FILING DATE: 2000-09-20
;   PRIOR APPLICATION NUMBER: US/60/234,923
;   PRIOR FILING DATE: 2000-09-25
;   PRIOR APPLICATION NUMBER: US/60/235,134
;   PRIOR FILING DATE: 2000-09-25
;   PRIOR APPLICATION NUMBER: US/60/235,637
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,638
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,711
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,720
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,840
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,863
;   PRIOR FILING DATE: 2000-09-27
;   NUMBER OF SEQ ID NOS: 2276
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1838
;   LENGTH: 2244
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-1838

Query Match             8.8%;  Score 16;  DB 10;  Length 2244;
Best Local Similarity  100.0%;  Pred. No. 71;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       112 AAAGTTGATGCTGAGA 127
             |||||||||||||||
Db       666 AAAGTTGATGCTGAGA 681

RESULT 25
US-09-925-301-415
; Sequence 415, Application US/09925301
; Patent No. US2002052308A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;   FILE REFERENCE: PA106
;   CURRENT APPLICATION NUMBER: US/09/925,301
;   CURRENT FILING DATE: 2001-08-10
;   PRIOR APPLICATION NUMBER: PCT/US00/05882
;   PRIOR FILING DATE: 2000-03-08
;   PRIOR APPLICATION NUMBER: 60/124,270
;   PRIOR FILING DATE: 1999-03-12
;   NUMBER OF SEQ ID NOS: 1694
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 415
;   LENGTH: 2880
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-415

Query Match             8.8%;  Score 16;  DB 10;  Length 2880;
Best Local Similarity  100.0%;  Pred. No. 73;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        57 TTATATGGACAATTGG 72
             |||||||||||||||
Db       418 TTATATGGACAATTGG 433

RESULT 26
US-09-851-129A-16
; Sequence 16, Application US/09851129A
; Patent No. US2002152486A1
; GENERAL INFORMATION:
;   APPLICANT: PAULIN, DENISE
;   APPLICANT: LI, ZHENLIN
;   APPLICANT: MERICSKAY, MATHIAS
;   TITLE OF INVENTION: VASCULAR SPECIFIC REGULATORY ELEMENTS CONTAINED IN THE
;   TITLE OF INVENTION: 5' FLANKING REGION
;   FILE REFERENCE: 205204US0
```

; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat...
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn version 3.0
SEQ ID NO 569
; LENGTH: 11990
; TYPE: DNA
; ORGANISM: Homosapiens
US-09-969-708-569

Query Match   8.8%;   Score 16;   DB 10;   Length 11990;
Best Local Similarity 100.0%;   Pred. No. 80;
Matches   16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   75 CCAGAACACAACAGCT 90
            |||||||||||||||||
Db   1390 CCAGAACACAACAGCT 1405

RESULT 29
US-10-138-618-35
; Sequence 35, Application US/10138618
; Publication No. US20030100525A1
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS CELLULAR APOPTOSIS, AND ITS MODULATION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/138,618
; FILING DATE: 06-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/975,080
; FILING DATE: 20-NOV-1997
; APPLICATION NUMBER: US 60/031,435
; FILING DATE: 20-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

; CURRENT APPLICATION NUMBER: US/09/851,129A
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 09/165,098
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 16
; LENGTH: 2889
; TYPE: DNA
; ORGANISM: Human
US-09-851-129A-16

Query Match   8.8%;   Score 16;   DB 10;   Length 2889;
Best Local Similarity 100.0%;   Pred. No. 73;
Matches   16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   75 CCAGAACACAACAGCT 90
            |||||||||||||||||
Db   1128 CCAGAACACAACAGCT 1143

RESULT 27
US-09-954-456-498/c
; Sequence 498, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc...
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
SEQ ID NO 498
; LENGTH: 6047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-498

Query Match   8.8%;   Score 16;   DB 10;   Length 6047;
Best Local Similarity 100.0%;   Pred. No. 76;
Matches   16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   17 CACAGACGCTGGAAGA 32
            |||||||||||||||||
Db   3355 CACAGACGCTGGAAGA 3340

RESULT 28
US-09-969-708-569
; Sequence 569, Application US/09969708
; Patent No. US20020102533A1

;   SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-10-138-618-35

Query Match   8.8%;   Score 16;   DB 9;   Length 14796;
Best Local Similarity   100.0%;   Pred. No. 81;
Matches 16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTTATTACTTATA 10824

RESULT 30
US-09-954-456-973
; Sequence 973, Application US/09954456
; Patent No. US2002011505761A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 973
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-973

Query Match   8.8%;   Score 16;   DB 10;   Length 14796;
Best Local Similarity   100.0%;   Pred. No. 81;
Matches 16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTTATTACTTATA 10824

RESULT 31
US-09-954-456-1636
; Sequence 1636, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1636
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1636

Query Match   8.8%;   Score 16;   DB 10;   Length 14796;
Best Local Similarity   100.0%;   Pred. No. 81;
Matches 16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   46 ATTTTATTACTTATA 61
          |||||||||||||||
Db 10809 ATTTTTATTACTTATA 10824

RESULT 32
US-09-918-186A-3
; Sequence 3, Application US/09918186A
; Patent No. US2002013770BA1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0585
; CURRENT APPLICATION NUMBER: US/09/918,186A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/496,694
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 250
; SEQ ID NO 3
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-09-918-186A-3

Query Match   8.8%;   Score 16;   DB 10;   Length 14796;
Best Local Similarity   100.0%;   Pred. No. 81;
Matches 16;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      46 ATTTTTATTACTTATA 61
           |||||||||||||||
Db   10809 ATTTTTATTACTTATA 10824

RESULT 33
US-09-880-107-3421
; Sequence 3421, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 4921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3421
;   LENGTH: 14796
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 U75285
US-09-880-107-3421

Query Match             8.8%;  Score 16;  DB 10;  Pred. No. 81;  Length 14796;
Best Local Similarity   100.0%;
Matches   16;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      46 ATTTTTATTACTTATA 61
           |||||||||||||||
Db   10809 ATTTTTATTACTTATA 10824

RESULT 34
US-09-764-877-3875/c
; Sequence 3875, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC005
;   CURRENT APPLICATION NUMBER: US/09/764,877
;   CURRENT FILING DATE: 2001-01-17
;   Prior application data removed - refer to PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 4031
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3875
;   LENGTH: 31314
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-3875

Query Match             8.8%;  Score 16;  DB 10;  Pred. No. 85;  Length 31314;
Best Local Similarity   100.0%;
Matches   16;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      42 AAGGATTTTTATTACT 57
           |||||||||||||||
Db   14284 AAGGATTTTTATTACT 14269

RESULT 35
US-10-175-523-71
; Sequence 71, Application US/10175523
```

```
; Publication No. US20030096264A1
; GENERAL INFORMATION:
;   APPLICANT: Brockman, Jeffrey
;   APPLICANT: Evans, David
;   APPLICANT: Hook, Derek
;   APPLICANT: Klimczak, Leszek
;   APPLICANT: Laeng, Pascal
;   APPLICANT: Palfreyman, Michael
;   APPLICANT: Rajan, Prithi
;   TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
;   FILE REFERENCE: 3235/1J795-US3
;   CURRENT APPLICATION NUMBER: US/10/175,523
;   CURRENT FILING DATE: 2002-06-18
;   PRIOR APPLICATION NUMBER: US 60/299,151
;   PRIOR FILING DATE: 2001-06-18
;   PRIOR APPLICATION NUMBER: US 60/317,828
;   PRIOR FILING DATE: 2001-09-07
;   PRIOR APPLICATION NUMBER: US 60/325,150
;   PRIOR FILING DATE: 2001-09-25
;   PRIOR APPLICATION NUMBER: US 60/333,047
;   PRIOR FILING DATE: 2001-11-14
;   PRIOR APPLICATION NUMBER: US 60/349,936
;   PRIOR FILING DATE: 2002-01-18
;   PRIOR APPLICATION NUMBER: US 60/361,834
;   PRIOR FILING DATE: 2002-03-04
;   NUMBER OF SEQ ID NOS: 197
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
;   LENGTH: 123192
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-175-523-71

Query Match             8.8%;  Score 16;  DB 9;  Pred. No. 92;  Length 123192;
Best Local Similarity   100.0%;
Matches   16;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      44 GGATTTTTATTACTTA 59
           |||||||||||||||
Db  118578 GGATTTTTATTACTTA 118593

RESULT 36
US-10-185-770-3
; Sequence 3, Application US/10185770
; Publication No. US20030022217A1
; GENERAL INFORMATION:
;   APPLICANT: CECCARDI, Toni et al.
;   TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;   TITLE OF INVENTION: USES THEREOF
;   FILE REFERENCE: CL0001247
;   CURRENT APPLICATION NUMBER: US/10/185,770
;   CURRENT FILING DATE: 2002-07-01
;   PRIOR APPLICATION NUMBER: 60/301,852
;   PRIOR FILING DATE: 2001-07-02
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 186957
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(186957)
;   OTHER INFORMATION: n = A,T,C or G
US-10-185-770-3

Query Match             8.8%;  Score 16;  DB 9;  Pred. No. 94;  Length 186957;
Best Local Similarity   100.0%;
Matches   16;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      54 TACTTATATGACACAT 69
```

```
Db     98266  |||||||||||||||||||||
               TACTTATATGGACAAT 98281
```

RESULT 37
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US2002012768TA1
; GENERAL INFORMATION:
;   APPLICANT: SHIGENOBU, SHUJI
;   APPLICANT: WATANABE, HIDEMI
;   APPLICANT: HATTORI, MASAHIRA
;   APPLICANT: SAKAKI, YOSHIYUKI
;   TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;   FILE REFERENCE: 081356/0159
;   CURRENT APPLICATION NUMBER: US/09/790,988
;   CURRENT FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: JP2000-107160
;   PRIOR FILING DATE: 2000-04-07
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match   8.8%; Score 16; DB 10; Length 640681;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   155 TGGTGATGATTGGAAT 170
           ||||||||||||||||
Db 324454 TGGTGATGATTGGAAT 324439

RESULT 38
US-09-263-959-1/c
; Sequence 1, Application US/09263959
; Patent No. US2002015089141
; GENERAL INFORMATION:
;   APPLICANT: Hood, Leroy E.
;   APPLICANT: Rowen, Lee
;   APPLICANT: Koop, Ben F.
;   TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;   NUMBER OF SEQUENCES: 1279
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Seed and Berry LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: US
;   ZIP: 98104-7092
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/263,959
;   FILING DATE: 05-MAR-1999
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: McMasters, David D.
;   REGISTRATION NUMBER: 33,963
;   REFERENCE/DOCKET NUMBER: 920010.426C2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 684973 base pairs
;   TYPE: nucleic acid
```

```
;     STRANDEDNESS: single
;     TOPOLOGY: linear
US-09-263-959-1
```

Query Match   8.8%; Score 16; DB 10; Length 684973;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   41 GAAGGATTTTATTAC 56
           ||||||||||||||||
Db 125801 GAAGGATTTTATTAC 125786

RESULT 39
US-09-864-761-23326
; Sequence 23326, Application US/09864761
; Patent No. US2002004876341
; GENERAL INFORMATION:
;   APPLICANT: Penn, Sharron G.
;   APPLICANT: Rank, David R.
;   APPLICANT: Hanzel, David K.
;   APPLICANT: Chen, Wensheng
;   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;   TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;   FILE REFERENCE: Aeomica-X-1
;   CURRENT APPLICATION NUMBER: US/09/864,761
;   CURRENT FILING DATE: 2001-05-23
;   PRIOR APPLICATION NUMBER: US 60/180,312
;   PRIOR FILING DATE: 2000-02-04
;   PRIOR APPLICATION NUMBER: US 60/207,456
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: US 09/632,366
;   PRIOR FILING DATE: 2000-08-03
;   PRIOR APPLICATION NUMBER: GB 24263.6
;   PRIOR FILING DATE: 2000-10-04
;   PRIOR APPLICATION NUMBER: US 60/236,359
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: PCT/US01/00666
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00667
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00664
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00669
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00665
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00668
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00663
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00662
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00661
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00670
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/234,687
;   PRIOR FILING DATE: 2000-09-21
;   PRIOR APPLICATION NUMBER: US 09/608,408
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: US 09/774,203
;   PRIOR FILING DATE: 2001-01-29
;   NUMBER OF SEQ ID NOS: 49117
;   SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23326
;   LENGTH: 81
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC018528.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EST HUMAN HIT: AI732364.1, EVALUE 9.70e-01
; OTHER INFORMATION: NT HIT: L49349.1, EVALUE 1.60e-01
US-09-864-761-23326

Query Match   8.3%; Score 15; DB 10; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   88 GCTGAGCAAGAGGCC 102
              ||||||||||||||
Db   28 GCTGAGCAAGAGGCC 42

RESULT 40
US-09-560-863-157
; Sequence 157, Application US/09560863
; Patent No. US2002011080 9A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael C.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US2002011080 9A1e1 Human Polynucleotides and the
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: LEX-0018-USA
; CURRENT APPLICATION NUMBER: US/09/560,863
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/132,408
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 133
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(133)
; OTHER INFORMATION: n = A,T,C or G
US-09-560-863-157

Query Match   8.3%; Score 15; DB 10; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   45 GATTTTTATTACTTA 59
              ||||||||||||||
Db   88 GATTTTTATTACTTA 102

RESULT 41
US-09-770-696-583/c
; Sequence 583, Application US/09770696
; Patent No. US2001004494 0A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2031US (PARA-020PRV)
; CURRENT APPLICATION NUMBER: US/09/770,696
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,278
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 583
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-696-583

Query Match   8.3%; Score 15; DB 10; Length 136;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   116 TTGATGCTGAGAACT 130
              ||||||||||||||
Db   106 TTGATGCTGAGAACT 92

RESULT 42
US-09-991-936-464/c
; Sequence 464, Application US/0991936
; Publication No. US20030073327A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisnewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 464
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-464

Query Match   8.3%; Score 15; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   107 AAGCCAAAGTTGATG 121
              ||||||||||||||
Db   119 AAGCCAAAGTTGATG 105

RESULT 43
US-09-864-761-20062
; Sequence 20062, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20062
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000193.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HeLA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: NT HIT: AL163280.2, EVALUE 4.00e-08
; OTHER INFORMATION: EST_HUMAN HIT: BE222239.1, EVALUE 3.00e-09
; OTHER INFORMATION: SWISSPROT HIT: P37246, EVALUE 2.50e+00
US-09-864-761-20062

Query Match   8.3%; Score 15; DB 10; Pred. No. 2.2e+02; Length 271;
Best Local Similarity   100.0%;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   15 CACACAGACGCTGGA 29
            |||||||||||||||
Db   14 CACACAGACGCTGGA 28

RESULT 44
US-09-878-574-11763
; Sequence 11763, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11763
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 70106486ZH1
US-09-878-574-11763

Query Match   8.3%; Score 15; DB 10; Length 271;
Best Local Similarity   100.0%; Pred. No. 2.2e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   108 AGCCAAGTTGATGC 122
            ||||||||||||||
Db   64 AGCCAAGTTGATGC 78

RESULT 45
US-09-864-761-19038
; Sequence 19038, Application US/09864761
; Patent No. US2002004876JA1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```
;        PRIOR APPLICATION NUMBER: PCT/US01/00661
;        PRIOR FILING DATE: 2001-01-30
;        PRIOR APPLICATION NUMBER: PCT/US01/00670
;        PRIOR FILING DATE: 2001-01-30
;        PRIOR APPLICATION NUMBER: US 60/234,687
;        PRIOR FILING DATE: 2000-09-21
;        PRIOR APPLICATION NUMBER: US 09/608,408
;        PRIOR FILING DATE: 2000-06-30
;        PRIOR APPLICATION NUMBER: US 09/774,203
;        PRIOR FILING DATE: 2001-01-29
;        NUMBER OF SEQ ID NOS: 49117
;        SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19038
;        LENGTH: 274
;        TYPE: DNA
;        ORGANISM: Homo sapiens
;        FEATURE:
;        OTHER INFORMATION: MAP TO AP000313.1
;        OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
;        OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.4
;        OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.1
;        OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
;        OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3
;        OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
;        OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8
;        OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.5
;        OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8
;        OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
;        OTHER INFORMATION: NT HIT: AL163280.2, EVALUE 4.00e-08
;        OTHER INFORMATION: EST HUMAN HIT: BE222239.1, EVALUE 3.00e-09
;        OTHER INFORMATION: SWISSPROT HIT: P37246, EVALUE 2.50e+00
US-09-864-761-19038

Query Match        8.3%;   Score 15;   DB 10;   Length 274;
Best Local Similarity   100.0%;   Pred. No. 2.2e+02;
Matches   15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       15 CACACAGACGCTGGA 29
            ||||||||||||||||
Db       14 CACACAGACGCTGGA 28

Search completed: June   9, 2003, 13:26:56
Job time : 283 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: June 9, 2003, 11:08:01 ; Search time 900 Seconds
                                    (without alignments)
                                    5852.898 Million cell updates/sec

Title: US-09-550-163-1_COPY_80_260
Perfect score: 181
Sequence: 1 acttatccaattcacaca........gattggaatgttcctcttca 181

Scoring table: OLIGO_NUC
          Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 1455140287 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*
        15: gb_ba:*
        16: em_fun:*
        17: em_hum:*
        18: em_in:*
        19: em_mu:*
        20: em_om:*
        21: em_or:*
        22: em_ov:*
        23: em_pat:*
        24: em_ph:*
        25: em_pl:*
        26: em_ro:*
        27: em_sts:*
        28: em_un:*
        29: em_vi:*
        30: em_htg_hum:*
        31: em_htg_inv:*
        32: em_htg_other:*
        33: em_htg_mus:*
        34: em_htg_pln:*
        35: em_htg_rod:*
        36: em_htg_mam:*
        37: em_htg_vrt:*
        38: em_sy:*
        39: em_htgo_hum:*
        40: em_htgo_mus:*
        41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	Str	ID	Description
1	181	100.0	732	6		AX406939	AX406939 Sequence
2	181	100.0	732	6		AX406941	AX406941 Sequence
3	181	100.0	732	9		AF071002	AF071002 Homo sapi
4	181	100.0	809	9		AF302095	AF302095 Homo sapi
5	181	100.0	24608	9		AP000320	AP000320 Homo sapi
6	181	100.0	100000	9		AP000052	AP000052 Homo sapi
7	181	100.0	100000	9		AP000167	AP000167 Homo sapi
8	181	100.0	100000	17		AP000120	Ap000120 Homo sapi
9	181	100.0	340000	6		AP001719	AP001719 Homo sapi
10	165	91.2	732	6		AX406947	AX406947 Sequence
11	163	90.1	732	6		AX406945	AX406945 Sequence
12	154	85.1	732	6		AX406943	AX406943 Sequence
13	26	14.4	372	10		AY050513	AY050513 Cavia por
14	26	14.4	468	10		AF071003	AF071003 Rattus no
15	26	14.4	1664	10		BC022699	BC022699 Mus muscu
16	26	14.4	144709	10	c	AC117904	AC117904 Rattus no
17	23	12.7	215	4		AF329636	AF329636 Oryctolag
18	23	12.7	225	4		AY079211	AY079211 Sus scrof
19	23	12.7	228	4		AF387764	AF387764 Equus cab
20	21	11.6	21	6		AX406951	AX406951 Sequence
21	20	11.0	20	6	c	AX406950	AX406950 Sequence
22	20	11.0	82479	2	c	AC108241	AC108241 Rattus no
23	20	11.0	169873	2	c	AC108558	AC108558 Rattus no
24	19	10.5	1665	9		AB050510	AB050510 Macaca fa
25	19	10.5	77129	8	c	AB023033	AB023033 Arabidops
26	19	10.5	100466	9		AC113931	AC113931 Homo sapi
27	19	10.5	109920	9		AL359851	AL359851 Human DNA
28	19	10.5	114771	9	c	HS569D19	AL022334 Human DNA
29	19	10.5	120531	2		AC097300	AC097300 Rattus no
30	19	10.5	129338	9		AC016597	AC016597 Homo sapi
31	19	10.5	164456	2		AC078804	AC078804 Rattus no
32	19	10.5	164995	2	c	AC098993	AC098993 Rattus no
33	19	10.5	170520	9		AC092130	AC092130 Homo sapi
34	19	10.5	176036	2		AC128107	AC128107 Rattus no
35	19	10.5	180153	2	c	AC127774	AC127774 Rattus no
36	19	10.5	183178	9		AC024134	AC024134 Homo sapi
37	19	10.5	185754	9		AC073162	AC073162 Homo sapi
38	19	10.5	199551	2	c	AC006281	AC006281 Plasmodiu
39	19	10.5	251199	9		HUAE000659	AE000659 Homo sapi
40	18	9.9	410	8	c	BRA417800	AJ417800 Brassica
41	18	9.9	659	11	c	G39242	G39242 Z20324 Zebr
42	18	9.9	750	9		HSUATM32	U55733 Human ataxi
43	18	9.9	996	9		AX414620	AX414620 Sequence
44	18	9.9	1841	1	c	PGHAGCG	Z27394 P.gingivali
45	18	9.9	1841	6		AR049926	AR049926 Sequence
46	18	9.9	1841	6		AR051685	AR051685 Sequence
47	18	9.9	2793	8		MZEPHD1	D73410 Zea mays mR
48	18	9.9	2804	6	c	AR005012	AR005012 Sequence
49	18	9.9	3505	1		U22019	MHU22019 Mycoplasma
50	18	9.9	3518	10	c	RNU76635	U76635 Rattus norv
51	18	9.9	3521	1		MHP120	X78450 M.hominis (
52	18	9.9	4425	9		OSA427978	AJ427978 Oryza sat
53	18	9.9	4985	6		AX416827	AX416827 Sequence
54	18	9.9	37219	9	c	HSN11D4	Z68285 Human DNA s
55	18	9.9	50000	6	c	AX392735	AX392735 Sequence
56	18	9.9	51440	2	c	AC129513	AC129513 Homo sapi
57	18	9.9	60932	2		AC105037	AC105037 Homo sapi
58	18	9.9	65582	2		AC130355	AC130355 Homo sapi
59	18	9.9	91370	6		AP003757	AP003757 Oryza sat
60	18	9.9	93499	2		AC093774	AC093774 Homo sapi
61	18	9.9	94038	2		AC111897	AC111897 Rattus no
62	18	9.9	103316	2	c	AC130439	AC130439 Homo sapi
63	18	9.9	107587	9		AL365255	AL365255 Human DNA
64	18	9.9	109666	9		AP000838	AP000838 Homo sapi
65	18	9.9	110000	2		AC026673_0	AC026673 Homo sapi

Pred. No. is the number of results predicted by chance to have a

#		Score	Len	Pos	N	Locus	Accession	Description
66	c	9.9	18	112518	2	AC097147	AC097147	Rattus no
67		9.9	18	114212	2	AL355856	AL355856	Homo sapi
68	c	9.9	18	116135	2	AC094899	AC094899	Rattus no
69	c	9.9	18	121362	9	AL359203	AL359203	Human DNA
70	c	9.9	18	122213	9	AC009507	AC009507	Homo sapi
71		9.9	18	124250	9	AC093387	AC093387	Homo sapi
72	c	9.9	18	135953	9	AL583854	AL583854	Human DNA
73		9.9	18	137955	2	AC109290	AC109290	Mus muscu
74		9.9	18	139629	2	AP003809	AP003809	Oryza sat
75		9.9	18	140539	2	AC079758	AC079758	Homo sapi
76	c	9.9	18	140788	9	HS800P24	Z99758	Human,DNA s
77		9.9	18	143063	9	AC040936	AC040936	Homo sapi
78		9.9	18	143848	2	AC067924	AC067924	Homo sapi
79		9.9	18	145603	9	AC005968	AC005968	Homo sapi
80	c	9.9	18	146158	9	AL136136	AL136136	Human DNA
81	c	9.9	18	146545	2	AC104009	AC104009	Homo sapi
82		9.9	18	151482	9	AC068765	AC068765	Homo sapi
83		9.9	18	156140	9	AC067745	AC067745	Homo sapi
84		9.9	18	161788	2	AC025639	AC025639	Homo sapi
85		9.9	18	161935	2	AC096075	AC096075	Rattus no
86	c	9.9	18	163535	2	AC026876	AC026876	Homo sapi
87		9.9	18	164431	2	AP003488	AP003488	Oryza sat
88		9.9	18	168663	9	AL354815	AL354815	Human DNA
89		9.9	18	168794	2	AC026252	AC026252	Homo sapi
90		9.9	18	169733	2	AC068046	AC068046	Homo sapi
91		9.9	18	170343	2	AP005185	AP005185	Oryza sat
92	c	9.9	18	170371	8	AB026295	AB026295	Oryza sat
93		9.9	18	172659	2	AC024320	AC024320	Homo sapi
94		9.9	18	174015	2	AC122664	AC122664	Rattus no
95		9.9	18	174024	9	AL583827	AL583827	Human chr
96		9.9	18	176399	2	CNS05TE1	AL355398	Human chr
97		9.9	18	180163	9	AP001925	AP001925	Homo sapi
98		9.9	18	180211	9	CNS05TDZ	AL358292	Human chr
99		9.9	18	180222	2	AC099224	AC099224	Rattus no
100		9.9	18	181977	2	AC103487	AC103487	Rattus no
101	c	9.9	18	184490	9	HSU82828	U82828	Homo sapien
102		9.9	18	186158	2	AC106303	AC106303	Rattus no
103		9.9	18	186215	9	AC097504	AC097504	Homo sapi
104		9.9	18	189476	2	AC009893	AC009893	Homo sapi
105		9.9	18	191179	2	AC091476	AC091476	Homo sapi
106		9.9	18	203959	9	AP001137	AP001137	Homo sapi
107		9.9	18	212038	2	AC110816	AC110816	Mus muscu
108		9.9	18	213462	10	AC077689	AC077689	Mus Muscu
109	c	9.9	18	215467	2	AC013420	AC013420	Homo sapi
110	c	9.9	18	215541	9	AC016904	AC016904	Homo sapi
111	c	9.9	18	218186	9	AC083865	AC083865	Homo sapi
112		9.9	18	219151	2	AC115353	AC115353	Rattus no
113		9.9	18	219769	9	AC067964	AC067964	Mus Muscu
114		9.9	18	229963	2	AC125174	AC125174	Mus muscu
115	c	9.4	17	232392	14	AF250284	AF250284	Ambacta m
116		9.4	17	239371	2	AC121602	AC121602	Mus muscu
117		9.4	17	254116	10	AL596136	AL596136	Mouse DNA
118	c	9.4	17	265469	2	AC103334	AC103334	Rattus no
119		9.4	17	280050	1	AL591975	AL591975	Listeria
120		9.4	17	340000	10	AP001681	AP001681	Homo sapi
121		9.4	17	343550	1	AP003587	AP003587	Nostoc sp
122		9.4	17	485	1	BSCOTDG	BSCOTDG	Bacillus su
123		9.4	17	676	2	HSA332056	AJ332056	Homo sapi
124		9.4	17	762	8	HSA331176	AJ331176	Triticum
125		9.4	17	790	8	AF378035	AF378035	Homo sapi
126		9.4	17	1293	10	AF415215	AF415215	Mus muscu
127		9.4	17	1422	10	AF099053	AF099053	Mus muscu
128		9.4	17	1754	8	AF274864	AF274864	Brassica
129	c	9.4	17	1859	10	BC009013	BC009013	Mus muscu
130	c	9.4	17	2147	10	AF247177	AF247177	Mus muscu
131	c	9.4	17	2199	10	AB004109	AB004109	Chinese h
132	c	9.4	17	2384	2	AK023309	AK023309	Homo sapi
133		9.4	17	2557	3	HVOCTOKS0	X95506	H.virescens
134	c	9.4	17	2693	6	AX073142	AX073142	Zea mays mR
135		9.4	17	2793	6	MZEPHD1		Sequence
136		9.4	17	2804	6	AR005012	AR005012	Sequence
137	c	9.4	17	3424	4	AF047484	AF047484	Vibrio vu
138	c	9.4	17	3641	8	SCYOR019W	Z74927	S.cerevisia
139	c	9.4	17	3725	9	AF133588	AF133588	Homo sapi
140		9.4	17	3822	9	AB023061	AB023061	Homo sapi
141		9.4	17	4016	3	BM6F2ECP	X77747	B.mori 6F6.
142		9.4	17	4056	14	AF389453	AF389453	Lymantria
143	c	9.4	17	8530	6	AX344411	AX344411	Sequence
144	c	9.4	17	8530	6	AX346657	AX346657	Sequence
145		9.4	17	9001	9	AC090020	AC090020	Homo sapi
146	c	9.4	17	23286	1	L47838	BACPONAYPP	Bacillus su
147		9.4	17	31607	2	AC013178	AC013178	Drosophil
148		9.4	17	41208	9	AL589906	AL589906	Human DNA
149		9.4	17	51565	9	AC007745	AC007745	Homo sapi
150		9.4	17	51770	5	AL590490	AL590490	Human DNA
151		9.4	17	57294	2	AC105049	AC105049	Homo sapi
152	c	9.4	17	57619	8	X87331	SCXV58KB	S.cerevisia
153		9.4	17	62370	2	AL451008	AL451008	Homo sapi
154	c	9.4	17	62377	2	AL360009	AL360009	Homo sapi
155		9.4	17	64677	2	AC079102	AC079102	Homo sapi
156	c	9.4	17	64931	10	AF320600	AF320600OS1	Mus muscu
157		9.4	17	66668	2	AC119961	AC119961	Mus muscu
158		9.4	17	68589	2	HS237J2	AL021394	Human DNA
159		9.4	17	70322	2	AC119219	AC119219	Mus muscu
160	c	9.4	17	70426	2	AC122618	AC122618	Rattus no
161	c	9.4	17	74109	8	NCB1A5	AL451109	Neurospor
162		9.4	17	77521	2	AC018562	AC018562	Homo sapi
163		9.4	17	83676	2	AC129856	AC129856	Rattus no
164		9.4	17	85910	2	AC098836	AC098836	Oryza sat
165	c	9.4	17	87172	3	AL772337_3	AL772337_3	Continuation (4 of
166		9.4	17	87705	9	HSDJ6315	AL049821	Human DNA
167		9.4	17	88509	9	HSJ159G19	AL078462	Human DNA
168		9.4	17	88823	2	AL713965	AL713965	Human DNA
169		9.4	17	89790	2	AC109412	AC109412	Rattus no
170		9.4	17	94244	2	AC063938	AC063938	Homo sapi
171		9.4	17	95736	9	AC037474	AC037474	Homo sapi
172		9.4	17	97386	2	AC037453	AC037453	Homo sapi
173	c	9.4	17	97961	9	AC098031	AC098031	Rattus no
174	c	9.4	17	98179	9	AC092363	AC092363	Homo sapi
175		9.4	17	99887	2	AC026481	AC026481	Homo sapi
176	c	9.4	17	100810	9	AC011991	AC011991	Homo sapi
177	c	9.4	17	101121	9	AC092919	AC092919	Homo sapi
178	c	9.4	17	102740	9	AC096785_3	AC096785_3	Rattus no
179	c	9.4	17	104215	2	AC123491	AC123491	Rattus no
180	c	9.4	17	104234	2	AC078907	AC078907	Homo sapi
181		9.4	17	105579	9	AF282855	AF282855	Homo sapi
182		9.4	17	109622	2	AC098598	AC098598	Oryza sat
183		9.4	17	110000	2	AC074222_2	AC074222_2	Continuation (3 of
184		9.4	17	110000	2	AC092450_0	AC092450_0	Homo sapi
185		9.4	17	110000	2	AC098564_3	AC098564_3	Continuation (4 of
186		9.4	17	110112	2	AC024907_0	AC024907_0	Homo sapi
187		9.4	17	111466	2	AC129090	AC129090	Homo sapi
188		9.4	17	111974	2	AC104274	AC104274	Oryza sat
189		9.4	17	112581	2	AC006156	AC006156	Homo sapi
190		9.4	17	112716	8	NCB8L3	AL513462	Neuropor
191		9.4	17	112735	2	AC126112	AC126112	Rattus no
192		9.4	17	112746	2	AC130995	AC130995	Rattus no
193		9.4	17	113000	2	AP005262	AP005262	Oryza sat
194		9.4	17	113677	2	AL513492	AL513492	Human DNA
195		9.4	17	115045	2	AC004414	AC004414	Homo sapi
196		9.4	17	115851	8	AC002505	AC002505	Arabidops
197	c	9.4	17	117352	2	AC116181	AC116181	Rattus no
198		9.4	17	121460	2	AC027330	AC027330	Homo sapi
199		9.4	17	121756	2	AC098671	AC098671	Homo sapi
200		9.4	17	123352	2	AC123516	AC123516	Oryza sat
201		9.4	17	123830	2	AC100244	AC100244	Mus muscu
202		9.4	17	124028	3	AC022417	AC022417	Homo sapi
203		9.4	17	125217	8	AC097627	AC097627	Genomic s
204		9.4	17	125990	2	AC108734	AC108734	Homo sapi
205		9.4	17	131331	2	AC108308	AC108308	Homo sapi
206		9.4	17	131607	3	HSB94K16	AL034429	Human DNA
207	c	9.4	17	131766	3	AC120451	AC120451	Rattus no
208	c	9.4	17	131974	2	HSDJ261K5	AL050350	Human DNA
209		9.4	17	132444	2	AC091878	AC091878	Homo sapi
210	c	9.4	17	132526	2	AC120946	AC120946	Rattus no
211	c	9.4	17	133894	2	AC105740	AC105740	Sus scrof

c	#	score	len	pos	n	accession	species
c	212	9.4	17	134060	9	AL358354	Human DNA
	213	9.4	17	134133	2	AC102587	Mus muscu
	214	9.4	17	134133	9	AC104320	Mus muscu
c	215	9.4	17	135055	2	AC011873	Homo sapi
	216	9.4	17	135091	2	AC107062	Homo sapi
	217	9.4	17	136115	2	AC120221	Rattus no
	218	9.4	17	136120	8	AC068923	Oryza sat
c	219	9.4	17	137391	2	AL161732	Human DNA
	220	9.4	17	138707	2	AC119710	Rattus no
	221	9.4	17	139086	9	AL356127	Human DNA
	222	9.4	17	139993	2	AC128283	Rattus no
	223	9.4	17	140582	2	AC128429	Rattus no
	224	9.4	17	140929	2	AC117805	Mus muscu
c	225	9.4	17	142690	2	AC102207	Homo sapi
	226	9.4	17	142979	2	AC000029	Homo sapi
	227	9.4	17	143220	9	AL389887	Human DNA
	228	9.4	17	143515	8	AP003252	Oryza sat
	229	9.4	17	144249	2	AC011262	Homo sapi
c	230	9.4	17	145784	9	AP001836	Homo sapi
	231	9.4	17	146395	2	AC127918	Rattus no
	232	9.4	17	146462	2	AC080142	Mus muscu
	233	9.4	17	146810	2	Z92540 / HS179115A	Human DNA s
	234	9.4	17	147085	2	AP005586	Oryza sat
	235	9.4	17	147437	2	AC118373	Rattus no
	236	9.4	17	148456	9	AL133550	Homo sapi
	237	9.4	17	149751	2	AC048339	Homo sapi
	238	9.4	17	149988	2	AC011929	Homo sapi
	239	9.4	17	151006	2	AC107255	Rattus no
	240	9.4	17	151418	2	AC118351	Rattus no
	241	9.4	17	152001	2	AC104569	Homo sapi
	242	9.4	17	152010	9	AC026785	Homo sapi
	243	9.4	17	152405	9	AF186190	Homo sapi
	244	9.4	17	152427	2	AC109368	Rattus no
	245	9.4	17	152552	9	AL356105	Homo sapi
	246	9.4	17	152602	2	AC026193	Homo sapi
c	247	9.4	17	153381	2	AP001990	Homo sapi
c	248	9.4	17	153607	9	AL355523	Homo sapi
	249	9.4	17	153805	9	AC090987	Homo sapi
c	250	9.4	17	155580	2	AC024597	Homo sapi
	251	9.4	17	155676	2	AC024946	Homo sapi
	252	9.4	17	155724	2	AC044780	Homo sapi
	253	9.4	17	156319	9	AC024714	Homo sapi
	254	9.4	17	157519	2	AC102038	Mus muscu
	255	9.4	17	157861	8	AP001389	Oryza sat
	256	9.4	17	158665	2	AC126902	Rattus no
	257	9.4	17	158665	9	AL356138	Homo sapi
	258	9.4	17	159198	2	AC115970	Mus muscu
	259	9.4	17	159871	2	AC124363	Rattus no
	260	9.4	17	160487	2	AC094402	Rattus no
	261	9.4	17	160779	10	AL732449	Mouse DNA
c	262	9.4	17	160851	2	AC105944	Mus muscu
	263	9.4	17	161417	9	AL356358	Human DNA
	264	9.4	17	161476	2	AC026321	Homo sapi
	265	9.4	17	161675	2	AP001975	Homo sapi
	266	9.4	17	161990	2	AC024723	Homo sapi
	267	9.4	17	162328	2	AC129300	Mus muscu
	268	9.4	17	162454	2	AC021568	Homo sapi
	269	9.4	17	162495	2	AC096084	Rattus no
c	270	9.4	17	162783	2	AC124391	Mus muscu
	271	9.4	17	162903	2	AC129144	Rattus no
	272	9.4	17	162925	2	AP004461	Oryza sat
	273	9.4	17	162943	2	AC090943	Homo sapi
	274	9.4	17	163027	9	AC002549	Homo sapi
	275	9.4	17	163731	2	AC008949	Homo sapi
	276	9.4	17	164396	2	AC095587	Rattus no
	277	9.4	17	164429	9	AC026366	Homo sapi
	278	9.4	17	165016	2	AC055742	Homo sapi
	279	9.4	17	165237	6	AX232505	Sequence
	280	9.4	17	165608	2	AC004492	Homo sapi
	281	9.4	17	166485	2	AC087878	Homo sapi
c	282	9.4	17	167212	9	AC127101	Homo sapi
	283	9.4	17	167227	2	AC025318	Rattus no
	284	9.4	17	167499	2	CNS01DWX / AL138975	Homo sapi

c	#	score	len	pos	n	accession	species
	285	9.4	17	168487	9	AL138820	Human DNA
	286	9.4	17	168528	9	AL355861	Human DNA
	287	9.4	17	169541	9	AL136178	Human DNA
c	288	9.4	17	169794	2	AC004688	Plasmodiu
c	289	9.4	17	170240	2	AL160279	Human DNA
	290	9.4	17	170283	9	AC122619	Rattus no
	291	9.4	17	170763	2	AC009196	Homo sapi
c	292	9.4	17	171656	2	AC102467	Homo sapi
c	293	9.4	17	171871	2	AC090768	Homo sapi
	294	9.4	17	172092	9	AC013341	Homo sapi
c	295	9.4	17	172380	2	AC095313	Rattus no
c	296	9.4	17	173317	2	AC016637	Homo sapi
c	297	9.4	17	173375	2	AC044815	Homo sapi
c	298	9.4	17	173414	2	AC125291	Drosophil
c	299	9.4	17	173712	2	AC098946	Homo sapi
c	300	9.4	17	174311	2	AC105632	Rattus no
	301	9.4	17	175132	2	AC008134	Homo sapi
	302	9.4	17	175144	2	AC068502	Mus muscu
	303	9.4	17	176291	2	AC024610	Homo sapi
	304	9.4	17	176877	2	AC126293	Rattus no
	305	9.4	17	176999	9	AC053518	Homo sapi
	306	9.4	17	177028	9	AF192304	Homo sapi
c	307	9.4	17	177108	2	AC026504	Homo sapi
c	308	9.4	17	177311	9	AP005467	Oryza sat
c	309	9.4	17	178033	2	AC107025	Homo sapi
	310	9.4	17	178069	2	AC006529	Homo sapi
c	311	9.4	17	178415	9	AC023547	Homo sapi
	312	9.4	17	178492	2	AC118405	Rattus no
	313	9.4	17	178857	2	AC113658	Rattus no
	314	9.4	17	179410	2	AC009698	Homo sapi
	315	9.4	17	179414	9	AC098933	Homo sapi
	316	9.4	17	180973	9	AC022316	Homo sapi
c	317	9.4	17	181016	9	AC068273	Homo sapi
c	318	9.4	17	181439	2	AC090611	Homo sapi
c	319	9.4	17	181739	2	AC116059	Rattus no
c	320	9.4	17	181837	9	AC062032	Homo sapi
c	321	9.4	17	182147	2	AC011190	Homo sapi
c	322	9.4	17	182496	9	AC023088	Homo sapi
c	323	9.4	17	182600	2	AC027527	Homo sapi
c	324	9.4	17	182624	2	AL807825	Mus muscu
	325	9.4	17	182714	2	AC127231	Mus muscu
c	326	9.4	17	184248	9	AC074199	Homo sapi
	327	9.4	17	184402	2	AC091291	Homo sapi
	328	9.4	17	184573	9	AC090740	Homo sapi
	329	9.4	17	184838	2	AC113710	Rattus no
	330	9.4	17	185648	2	AC105460	Homo sapi
	331	9.4	17	185755	9	AC115262	Homo sapi
	332	9.4	17	186184	2	AC020573	Homo sapi
	333	9.4	17	186496	2	AC078844	Homo sapi
	334	9.4	17	186927	9	AC122325	Mus muscu
	335	9.4	17	187136	9	AC055822	Homo sapi
	336	9.4	17	187833	2	AC127260	Mus muscu
	337	9.4	17	188096	2	AC102068	Mus muscu
	338	9.4	17	188224	2	AC007528	Homo sapi
	339	9.4	17	188451	2	AC022682	Mus muscu
	340	9.4	17	189129	9	AC112062	Homo sapi
	341	9.4	17	189694	2	AC090430	Mus muscu
	342	9.4	17	189974	2	AC020573	Rattus no
	343	9.4	17	190444	2	AC098121	Rattus no
	344	9.4	17	190868	2	AC011710	Homo sapi
c	345	9.4	17	191866	9	AC068676	Homo sapi
c	346	9.4	17	192531	9	AL772193	Zebrafish
c	347	9.4	17	192658	2	AC116702	Mus muscu
c	348	9.4	17	192843	10	AL607124	Mouse DNA
c	349	9.4	17	192952	2	AC022268	Homo sapi
c	350	9.4	17	193975	9	AC109970	Rattus no
c	351	9.4	17	194057	2	AC116940	Pan trogl
c	352	9.4	17	195229	2	AC124545	Mus muscu
c	353	9.4	17	196085	6	AC016737	Homo sapi
c	354	9.4	17	196516	2	AL844218	Mus muscu
c	355	9.4	17	197550	2	AL162171	Human chr
c	356	9.4	17	197611	9	AL390840	Human DNA
	357	9.4	17	198233	2	AC130067	Rattus no

No.	Score	Length	Position	Accession	N	Description
358	9.4	17	198306	AL837510	2	AL837510 Mus muscu
359	9.4	17	198422	AC113939	2	AC113939 Mus muscu
c 360	9.4	17	198901	AC116303	2	AC116303 Homo sapi
361	9.4	17	200311	AC123769	2	AC123769 Homo sapi
c 362	9.4	17	201726	AC025552	2	AC025552 Homo sapi
363	9.4	17	202111	AC096784	10	AC096784 Mus muscu
364	9.4	17	202163	AF235099	2	AF235099 Homo sapi
365	9.4	17	202655	AC087276	2	AC087276 Homo sapi
366	9.4	17	202763	AC005181	2	AC005181 Homo sapi
367	9.4	17	202866	AC092109	2	AC092109 Homo sapi
368	9.4	17	203395	AC073596	2	AC073596 Homo sapi
369	9.4	17	203579	AL645762	2	AL645762 Homo sapi
370	9.4	17	208307	AC073351	2	AC073351 Mus muscu
371	9.4	17	208342	AC105077	2	AC105077 Mus muscu
c 372	9.4	17	209636	AC121084	2	AC121084 Homo sapi
373	9.4	17	210000	AC009527	2	AC009527 Homo sapi
c 374	9.4	17	210361	AL671335	2	AL671335 Mus muscu
375	9.4	17	210672	HS179115	2	Z84464 Homo sapien
376	9.4	17	216707	AC074305	2	AC074305 Mus muscu
377	9.4	17	211759	AC106937	2	AC106937 Rattus no
378	9.4	17	212638	AC114920	2	AC114920 Mus muscu
379	9.4	17	212681	AC107701	2	AC107701 Mus muscu
c 380	9.4	17	212884	AC124443	2	AC124443 Mus muscu
381	9.4	17	213296	AC067784	2	AC067784 Homo sapi
382	9.4	17	213308	AC098281	2	AC098281 Rattus no
383	9.4	17	213680	BSUB0012	2	299115 Bacillus su
c 384	9.4	17	214000	AC093317	10	AC093317 Mus muscu
385	9.4	17	215260	AC009600	2	AC009600 Homo sapi
386	9.4	17	215308	AC114967	9	AC114967 Homo sapi
387	9.4	17	216733	AL583892	10	AL583892 Mouse DNA
388	9.4	17	216825	AC108908	2	AC108908 Mus muscu
389	9.4	17	217250	AC018595	2	AC018595 Homo sapi
c 390	9.4	17	217286	AC098713	10	AC098713 Mus muscu
391	9.4	17	217290	AC074332	2	AC074332 Mus muscu
392	9.4	17	220546	AC124742	2	AC124742 Mus muscu
393	9.4	17	220623	AE003451	3	AE003451 Drosophil
c 394	9.4	17	222310	AC073589	10	AC073589 Mus muscu
c 395	9.4	17	222437	AC121849	2	AC121849 Mus muscu
396	9.4	17	222806	AC018977	2	AC018977 Homo sapi
397	9.4	17	224299	AC130533	2	AC130533 Mus muscu
398	9.4	17	224406	AC098244	2	AC098244 Rattus no
399	9.4	17	225518	AC126458	2	AC126458 Mus muscu
400	9.4	17	227028	AL808140	2	AL808140 Mus muscu
401	9.4	17	230595	AC091745	2	AC091745 Homo sapi
402	9.4	17	233867	AC125541	2	AC125541 Mus muscu
403	9.4	17	242918	AC127347	2	AC127347 Mus muscu
404	9.4	17	243179	AC123863	9	AC123863 Mus muscu
405	9.4	17	251124	HUAE000660	2	AE000660 Homo sapi
406	9.4	17	252358	AC126686	9	AC126686 Mus muscu
c 407	9.4	17	257912	AC126280	2	AC126280 Mus muscu
c 408	9.4	17	263649	AC125037	2	AC125037 Mus muscu
c 409	9.4	17	264778	AC122941	2	AC122941 Mus muscu
c 410	9.4	17	275253	AC068772	2	AC068772 Homo sapi
411	9.4	17	282375	AC101820	2	AC101820 Mus muscu
412	9.4	17	296683	AC124346	2	AC124346 Mus muscu
c 413	9.4	17	301450	AP003185	1	AP003185 Closttridi
414	9.4	17	306130	AC110996	2	AC110996 Homo sapi
415	9.4	17	308676	AC068639	2	AC068639 Homo sapi
416	8.8	16	165	G65099	11	G65099 stb13sl.t7
c 417	8.8	16	326	AF529292	11	AF529292 Cookeina
c 418	8.8	16	334	G10245	11	G10245 human STS C
c 419	8.8	16	343	AF529291	8	AF529291 Cookeina
c 420	8.8	16	379	G51759	11	G51759 SHGC-84529
421	8.8	16	385	AF151796	3	AF151796 Buthus ma
422	8.8	16	394	G09033	11	G09033 human STS C
423	8.8	16	559	AF394005	8	AF394005 Cookeina
424	8.8	16	601	AF394009	8	AF394009 Cookeina
c 425	8.8	16	610	AF394010	8	AF394010 Cookeina
c 426	8.8	16	631	AF394013	8	AF394013 Cookeina
427	8.8	16	651	AF393361	9	AF393361 Homo sapi
c 428	8.8	16	656	AF394012	8	AF394012 Cookeina
c 429	8.8	16	674	AF394007	8	AF394007 Cookeina
c 430	8.8	16	677	AF394006	8	AF394006 Cookeina
c 431	8.8	16	683	AF394008	8	AF394008 Cookeina
c 432	8.8	16	689	AF394011	8	AF394011 Cookeina
433	8.8	16	752	A09290	6	A09290 Human papil
434	8.8	16	758	HSA341865	9	AJ341865 Homo sapi
c 435	8.8	16	780	CNS01DFU	8	AL116770 Botrytis
436	8.8	16	844	CNS06EWD	11	AL395651 T7 end of
437	8.8	16	930	AF282285	10	AF282285 Mus muscu
438	8.8	16	930	AY073615	10	AY073615 Mus muscu
439	8.8	16	937	AX241839	6	AX241839 Sequence
440	8.8	16	996	AF195465	7	AF195465 Pseudoalt
441	8.8	16	1053	AF438139	3	AF438139 Styphlome
442	8.8	16	1074	AX241830	3	AX241830 Sequence
443	8.8	16	1078	AF398638	3	AF398638 Aptinus d
c 444	8.8	16	1078	AF398710	3	AF398710 Aptinus a
445	8.8	16	1092	AF514996	9	AF514996 Homo sapi
446	8.8	16	1093	AB041154S3	3	AB041356 Homo sapi
c 447	8.8	16	1094	AF398641	3	AF398641 Styphlodr
448	8.8	16	1453	AF483206	3	AF483206 Planococc
449	8.8	16	1520	D86227	9	D86227 Homo sapien
450	8.8	16	1520	D86228	9	D86228 Human clone
451	8.8	16	1520	D86230	9	D86230 Homo sapien
452	8.8	16	1521	I02377	6	I02377 Sequence 2
453	8.8	16	1549	HSM800602	8	AL080099 Homo sapi
454	8.8	16	1552	AB028247	14	AB028247 Norwalk-1
455	8.8	16	1593	BD011872	6	BD011872 Detection
456	8.8	16	1593	AB031013	14	AB031013 Norwalk-1
457	8.8	16	1598	BD004352	23	Bd004352 Detection
458	8.8	16	1603	HUMDMPSX	9	M36661 Human oroti
459	8.8	16	1698	STYGDHA	1	M24021 S.typhimuri
460	8.8	16	1711	AB062285	9	AB062285 Homo sapi
461	8.8	16	1711	BC000064	9	BC000064 Homo sapi
462	8.8	16	1750	BC007511	9	BC007511 Homo sapi
c 463	8.8	16	1809	AY087912	9	AY087912 Arabidops
464	8.8	16	1818	AF431970	9	AF431970 Homo sapi
c 465	8.8	16	1837	AF267223	1	AF267223 Candidatu
466	8.8	16	1902	AK056940	6	AK056940 Homo sapi
467	8.8	16	2042	AX077616	6	AX077616 Sequence
468	8.8	16	2045	BTU77588	4	U77588 Bos taurus
469	8.8	16	2093	D89631	4	D89631 Arabidopsis
470	8.8	16	2166	AF285099	3	AF285099 Ovis arie
471	8.8	16	2228	BC032942	4	BC032942 Homo sapi
472	8.8	16	2244	AK023335	6	AK023335 Homo sapi
473	8.8	16	2244	AX333692	4	AX333692 Sequence
474	8.8	16	2244	AX334356	6	AX334356 Sequence
475	8.8	16	2259	HUMUMPS	4	J03626 Human UMP s
476	8.8	16	2263	AF281677	9	AF281677 Bos tauru
477	8.8	16	2266	LMU29952	3	U29952 Leishmania
478	8.8	16	2315	AY099873	5	AY099873 Arabidops
479	8.8	16	2326	DYGSV2TRAN	5	L23403 Discopyge o
480	8.8	16	2371	A14690	1	A14690 HPV-1a DNA
481	8.8	16	2501	U01165	5	U01165 Caulobacter
482	8.8	16	2516	OMEREX1A	14	NOR89JB Norwalk vir
483	8.8	16	2558	AY071327	9	AY071327 Drosophil
484	8.8	16	2684	AK055329	10	MMU250723 Mus muscu
c 485	8.8	16	2942	HSDESMIN	8	X55154 Human gene
486	8.8	16	3284	AF305912	3	AF305912 Hordeum v
487	8.8	16	3338	AY051863	9	AY051863 Drosophil
488	8.8	16	3423	AF255625	9	AF255625 Homo sapi
489	8.8	16	3431	AF136976	9	AF136976 Homo sapi
490	8.8	16	3492	BC012430	10	BC012430 Homo sapi
491	8.8	16	3497	MMU250724	3	M14778 D.melanogas
492	8.8	16	3569	DROINSLR	9	AF016991 Callithri
493	8.8	16	3586	AF016991	2	AF321383S2 Callithri
494	8.8	16	3605	AF321383S2	2	AC017474 Drosophil
c 495	8.8	16	3607	AC017474	1	AE006709 Sulfolobu
c 496	8.8	16	4930	AE006709	1	TBU11682 Trypanoplas
c 497	8.8	16	5167	TBU11682	9	HS200KDU5 H.sapiens g
498	8.8	16	5293	HS200KDU5	6	270200 H.sapiens g
499	8.8	16	5350			AX346057 Sequence
c 500	8.8	16	5807	AX346057	6	AX346057 Sequence
c 501	8.8	16	5962	AX346188	6	AX346188 Sequence
c 502	8.8	16	6047	AX333016	6	AX333016 Sequence
c 503	8.8	16				

Left block

Mk	Idx			Num			Accession / Description	Code	Ct
c	504	8.8	16	6047	8.8	16	D86979 Homo sapien	D86979	9
c	505	8.8	16	6315	8.8	16	Z72902 S.cerevisia	SCYGR117C	9
c	506	8.8	16	6376	8.8	16	AJ242867 Anabaena	APC242867	1
c	507	8.8	16	6413	8.8	16	AF067776 Abiotroph	AF067776	1
	508	8.8	16	6463	8.8	16	U18351 Drosophila	DMU18351	3
	509	8.8	16	6757	8.8	16	AF450140 Mus muscu	AF450138S3	10
	510	8.8	16	7012	8.8	16	AB012048 Arabidops	AB012048	3
	511	8.8	16	7160	8.8	16	U28136 Drosophila	U28136	8
	512	8.8	16	7524	8.8	16	AF285100 Bos tauru	AF285100	4
c	513	8.8	16	7654	8.8	16	CVXRNA	CVXRNA	14
	514	8.8	16	7801	8.8	16	HPU06714 Human papil	HPU06714	14
	515	8.8	16	7815	8.8	16	V01116 Human papil	PAPAPI	14
	516	8.8	16	8649	8.8	16	AE005931 Caulobact	AE005931	1
c	517	8.8	16	8803	8.8	16	AE005807 Caulobact	AE005807	8
	518	8.8	16	9364	8.8	16	AF098972 Cryptococ	AF098972	1
	519	8.8	16	9383	8.8	16	AC020531 Drosophil	AC020531	1
	520	8.8	16	9870	8.8	16	AE011902 Xanthomon	AE011902	1
c	521	8.8	16	10432	8.8	16	AE011103 Methanosa	AE011103	1
	522	8.8	16	10620	8.8	16	AF052602 Danio rer	AF052602	5
	523	8.8	16	10692	8.8	16	AE010244 Pyrococcu	AE010244	1
	524	8.8	16	10988	8.8	16	AL592286 Human DNA	AL592286	9
	525	8.8	16	11252	8.8	16	AE005760 Caulobact	AE005760	1
	526	8.8	16	11990	8.8	16	AX329652 Sequence	AX329652	6
	527	8.8	16	11990	8.8	16	AX337531 Sequence	AX337531	6
	528	8.8	16	11990	8.8	16	M63391 Homo sapien	HUMDES	9
	529	8.8	16	12251	8.8	16	AE013897 Yersinia	AE013897	1
	530	8.8	16	13423	8.8	16	AE006416 Lactococc	AE006416	1
	531	8.8	16	14026	8.8	16	AC019852 Drosophil	AC019852	1
	532	8.8	16	14477	8.8	16	AE005610 Escherich	AE005610	1
	533	8.8	16	14529	8.8	16	AE001397 Plasmodiu	AE001397	3
	534	8.8	16	14745	8.8	16	AF143314 Homo sapi	PTEN3	9
	535	8.8	16	14796	8.8	16	AR157507 Sequence	AR157507	9
	536	8.8	16	14796	8.8	16	AR181541 Sequence	AR181541	9
	537	8.8	16	14796	8.8	16	AX333491 Sequence	AX333491	6
	538	8.8	16	14796	8.8	16	AX334154 Sequence	AX334154	6
	539	8.8	16	14796	8.8	16	AX336483 Sequence	AX336483	6
	540	8.8	16	14796	8.8	16	AX410775 Sequence	AX410775	6
	541	8.8	16	14796	8.8	16	U75285 Homo sapien	HSU75285	9
	542	8.8	16	17194	8.8	16	AE003000 Drosophil	AE003000	3
	543	8.8	16	17406	8.8	16	AF043433 Anopheles	AF043433	1
	544	8.8	16	18292	8.8	16	U21600 Streptomy	SCF76	1
	545	8.8	16	24523	8.8	16	Z50045 Caenorhabdi	CEF38B2	1
	546	8.8	16	29977	8.8	16	AE008756 Salmonell	AE008756	1
c	547	8.8	16	30163	8.8	16	AL158834 Human DNA	AL158834	9
	548	8.8	16	30299	8.8	16	AC110382 Mus muscu	AC110382	2
	549	8.8	16	30415	8.8	16	AE001577 Borrelia	AE001577	1
c	550	8.8	16	30794	8.8	16	AF016441 Caenorhab	AF016441	1
	551	8.8	16	30911	8.8	16	AC020208 Drosophil	AC020208	3
	552	8.8	16	31288	8.8	16	AC127872 Rattus no	AC127872	2
	553	8.8	16	31574	8.8	16	AF390900 Mus muscu	AF390897S4	10
	554	8.8	16	32668	8.8	16	AJ314858 Mus muscu	MMU314858	2
c	555	8.8	16	33839	8.8	16	AC014113 Drosophil	AC014113	3
	556	8.8	16	34046	8.8	16	U00035 Caenorhabdi	U00035	1
	557	8.8	16	34089	8.8	16	AC013149 Drosophil	AC013149	3
	558	8.8	16	35438	8.8	16	AC020245 Drosophil	AC020245	3
	559	8.8	16	37262	8.8	16	AC011527 Homo sapi	AC011527	9
c	560	8.8	16	38055	8.8	16	AP001299 Arabidops	AP001299	8
	561	8.8	16	38629	8.8	16	AC006190 Homo sapi	AC006190	9
	562	8.8	16	40036	8.8	16	AL590709 Human DNA	AL590709	9
	563	8.8	16	40697	8.8	16	AC101084 Mus muscu	AC101084	2
c	564	8.8	16	40790	8.8	16	AC107874 Homo sapi	AC107874	9
	565	8.8	16	41565	8.8	16	Z82264 Caenorhab	CEC4C3	1
	566	8.8	16	42688	8.8	16	AL032647 Caenorhab	CEY57A10B	1
	567	8.8	16	44727	8.8	16	AC084506 Caenorhab	CBRG15A11	1
	568	8.8	16	45744	8.8	16	AC010500 Homo sapi	AC010500	9
	569	8.8	16	45911	8.8	16	AC091413 Rattus no	AC091413	2
	570	8.8	16	46105	8.8	16	AB017070 Arabidops	AB017070	8
	571	8.8	16	46990	8.8	16	AC021085 Homo sapi	AC021085	9
c	572	8.8	16	49852	8.8	16	AC112645 Homo sapi	AC112645	9
c	573	8.8	16	51313	8.8	16	AL353587 Human DNA	AL353587	9
	574	8.8	16	51900	8.8	16	AC107005 Rattus no	AC107005	2
	575	8.8	16	51920	8.8	16	AC130226 Rattus no	AC130226	2
c	576	8.8	16	53524	8.8	16	AC098371 Rattus no	AC098371	2

Right block

Mk	Idx			Num			Code	Ct	Accession / Description
c	577	8.8	16	53561	8.8	16	AE000790	1	AE000790 Borrelia
c	578	8.8	16	53614	8.8	16	AC018481	2	AC018481 Drosophil
c	579	8.8	16	53749	8.8	16	AC091686	2	AC091686 Homo sapi
c	580	8.8	16	55113	8.8	16	HSDJ249H1	9	AL132660 Human DNA
c	581	8.8	16	56423	8.8	16	AB065501	9	AB065501 Homo sapi
	582	8.8	16	58527	8.8	16	AB007649	8	AB007649 Arabidops
	583	8.8	16	58785	8.8	16	AC114624	2	AC114624 Mus muscu
c	584	8.8	16	60000	8.8	16	AF049236	8	AF049236 Arabidops
	585	8.8	16	60085	8.8	16	AC111079	2	AC111079 Mus muscu
	586	8.8	16	61017	8.8	16	AC099990	2	AC099990 Mus muscu
	587	8.8	16	61119	8.8	16	AC074369	2	AC074369 Homo sapi
	588	8.8	16	61611	8.8	16	AL669936	10	AL669936 Mouse DNA
c	589	8.8	16	61340	8.8	16	AC093652	9	AC093652 Homo sapi
	590	8.8	16	61439	8.8	16	AC114416	2	AC114416 Mus muscu
	591	8.8	16	62305	8.8	16	AC090546	2	AC090546 Homo sapi
	592	8.8	16	62305	8.8	16	AC090546	2	AC090546 Homo sapi
c	593	8.8	16	62670	8.8	16	AC010566	3	AC010566 Drosophil
	594	8.8	16	63433	8.8	16	AC004320	3	AC004320 Drosophil
c	595	8.8	16	63731	8.8	16	AL360016 [4]	2	Continuation (5 of
c	596	8.8	16	64064	8.8	16	AC117767	2	AC117767 Mus muscu
c	597	8.8	16	64065	8.8	16	AC124253	9	AC124253 Human DNA
c	598	8.8	16	65971	8.8	16	AC013231	2	AC013231 Drosophil
c	599	8.8	16	66286	8.8	16	AC123765	2	AC123765 Homo sapi
	600	8.8	16	67077	8.8	16	AC100172	2	AC100172 Mus muscu
c	601	8.8	16	67246	8.8	16	AC087523	2	AC087523 Homo sapi
	602	8.8	16	67669	8.8	16	AC123780	2	AC123780 Homo sapi
c	603	8.8	16	68040	8.8	16	AP002082	9	AP002082 Homo sapi
	604	8.8	16	68171	8.8	16	AL158205	9	AL158205 Human DNA
	605	8.8	16	68462	8.8	16	AC044854	2	AC044854 Homo sapi
	606	8.8	16	68553	8.8	16	AC125030	2	AC125030 Mus muscu
c	607	8.8	16	68553	8.8	16	AC125030	2	AC125030 Mus muscu
	608	8.8	16	68584	8.8	16	AC019069	9	AC019069 Homo sapi
	609	8.8	16	68764	8.8	16	AL389921	9	AL389921 Human DNA
c	610	8.8	16	69402	8.8	16	CNS08CB8	9	AL845346 Oryza sat
	611	8.8	16	69670	8.8	16	AC114988	2	AC114988 Mus muscu
	612	8.8	16	70027	8.8	16	AC113454	2	AC113454 Homo sapi
c	613	8.8	16	70287	8.8	16	AC105128	2	AC105128 Homo sapi
c	614	8.8	16	70627	8.8	16	AC102436	2	AC102436 Mus muscu
	615	8.8	16	71118	8.8	16	AP000586	9	AP000586 Homo sapi
	616	8.8	16	71190	8.8	16	AC115634	9	AC115634 Oryza sat
	617	8.8	16	71257	8.8	16	AL590868	9	AL590868 Human DNA
c	618	8.8	16	71538	8.8	16	AC118243	2	AC118243 Mus muscu
	619	8.8	16	72354	8.8	16	AC130476	2	AC130476 Mus muscu
	620	8.8	16	72558	8.8	16	AC067767	9	AC067767 Homo sapi
	621	8.8	16	73754	8.8	16	AP000592	9	AP000592 Homo sapi
	622	8.8	16	74289	8.8	16	AC011813	9	AC011813 Homo sapi
	623	8.8	16	74610	8.8	16	AC114223	2	AC114223 Rattus no
c	624	8.8	16	74681	8.8	16	AC112904	2	AC112904 Drosophil
c	625	8.8	16	74681	8.8	16	AC020320	3	AC020320 Drosophil
	626	8.8	16	76477	8.8	16	AC022977	9	AC022977 Homo sapi
	627	8.8	16	77319	8.8	16	AC110757	9	AC110757 Homo sapi
	628	8.8	16	78605	8.8	16	AC113843	2	AC113843 Rattus no
	629	8.8	16	79044	8.8	16	AC021326	2	AC021326 Homo sapi
	630	8.8	16	79680	8.8	16	CNS08CA0	9	AL772422 Oryza sat
c	631	8.8	16	79953	8.8	16	AC097842	2	AC097842 Rattus no
	632	8.8	16	81670	8.8	16	AC012061	9	AC012061 Homo sapi
	633	8.8	16	83220	8.8	16	AC021501	2	AC021501 Rattus no
	634	8.8	16	83463	8.8	16	AL354874	9	AL354874 Human DNA
c	635	8.8	16	85718	8.8	16	AL356134	9	AL356134 Human DNA
	636	8.8	16	85718	8.8	16	AP005607	9	AP005607 Oryza sat
c	637	8.8	16	87246	8.8	16	AC008438	9	AC008438 Homo sapi
	638	8.8	16	87967	8.8	16	AL050329	9	AL050329 Human DNA
	639	8.8	16	88064	8.8	16	AC002539	9	AC002539 Homo sapi
	640	8.8	16	88223	8.8	16	AC005223	8	AC005223 Arabidops
	641	8.8	16	88621	8.8	16	AC022707	2	AC022707 Homo sapi
	642	8.8	16	88626	8.8	16	AC084755	9	AC084755 Homo sapi
	643	8.8	16	88826	8.8	16	AP004128	9	AP004128 Oryza sat
	644	8.8	16	89703	8.8	16	AC021501	2	AC021501 Rattus no
	645	8.8	16	89938	8.8	16	AC122973	2	AC122973 Rattus no
	646	8.8	16	90736	8.8	16	AL161644	9	AL161644 Human DNA
c	647	8.8	16	90823	8.8	16	HS299D3	8	Z84468 Human DNA s
	648	8.8	16	90854	8.8	16	ATT209	8	AL138658 Arabidops
c	649	8.8	16	91350	8.8	16	HSJ586J11	9	AL109828 Human DNA

Left group (rows 650–722)

	#	Name			Pos	Cnt	Accession	Species
	650	AC106810	8.8	16	91853	9	AC106810	Homo sapi
c	651	AC010393	8.8	16	92819	2	AC010393	Homo sapi
c	652	AC110010_3	8.8	16	93432	9	Continuation (4 of	
	653	AL359971	8.8	16	94019	2	AL359971	Human DNA
c	654	AC091231	8.8	16	95727	9	AC091231	Homo sapi
c	655	AC025775	8.8	16	95782	9	AC025775	Homo sapi
	656	AC097756	8.8	16	96026	2	AC097756	Rattus no
	657	AC093669	8.8	16	96120	9	AC093669	Homo sapi
	658	AL731575	8.8	16	96212	2	AL731575	Human DNA
c	659	AC096257	8.8	16	97165	2	AC096257	Rattus no
c	660	AP004471	8.8	16	97710	8	AP004471	Lotus jap
	661	AC091816	8.8	16	97860	9	AC091816	Homo sapi
c	662	AC105596	8.8	16	97970	2	AC105596	Rattus no
	663	AP002907	8.8	16	98146	9	AP002907	Homo sapi
	664	AC111150	8.8	16	98856	9	AC111150	Rattus no
	665	AC123261	8.8	16	99013	2	AC123261	Rattus no
	666	AL391361	8.8	16	99162	9	AL391361	Human DNA
	667	AC104637	8.8	16	99960	2	AC104637	Homo sapi
	668	AC096298	8.8	16	100282	9	AC096298	Rattus no
	669	AC026436	8.8	16	100899	9	AC026436	Homo sapi
	670	AC118492	8.8	16	101149	2	AC118492	Rattus no
	671	AC004038	8.8	16	101615	9	AC004038	Homo sapi
c	672	AC026735	8.8	16	102255	2	AC026735	Homo sapi
c	673	AC113707	8.8	16	102659	2	AC113707	Rattus no
	674	AC099683	8.8	16	102867	3	AC099683	Caenorhab
	675	AC095774	8.8	16	102871	2	AC095774	Rattus no
	676	AC034223	8.8	16	102905	9	AC034223	Homo sapi
	677	AC004893	8.8	16	103738	9	AC004893	Homo sapi
	678	ATT8B10	8.8	16	103787	8	AL138646	Arabidops
	679	AC119635	8.8	16	103894	2	AC119635	Rattus no
	680	AC107401	8.8	16	104079	9	AC107401	Homo sapi
	681	AC128414	8.8	16	104193	2	AC128414	Rattus no
	682	AC108064	8.8	16	104623	2	AC108064	Homo sapi
	683	AC117882	8.8	16	104627	2	AC117882	Rattus no
	684	AC123102	8.8	16	105021	2	AC123102	Rattus no
	685	AC110076	8.8	16	105211	2	AC110076	Homo sapi
	686	AC113808	8.8	16	105922	2	AC113808	Rattus no
c	687	AC068792	8.8	16	106277	9	AC068792	Homo sapi
c	688	AC080103	8.8	16	106444	2	AC080103	Rattus no
	689	AC126312	8.8	16	106684	2	AC126312	Rattus no
	690	AC105516	8.8	16	107462	9	AC105516	Rattus no
	691	HS29H4	8.8	16	107469	2	AL078475	Human DNA
c	692	AC095639	8.8	16	108611	2	AC095639	Rattus no
	693	CNS08C8F	8.8	16	108805	2	AL731875	Oryza sat
c	694	AC096016	8.8	16	108975	2	AC096016	Rattus no
	695	AP000681	8.8	16	109149	9	AP000681	Homo sapi
	696	AC002328	8.8	16	109171	8	AC002328	Genomic s
	697	AL160031	8.8	16	109391	9	AL160031	Human DNA
	698	AC004535	8.8	16	109469	9	AC004535	Homo sapi
	699	AC097442	8.8	16	109552	2	AC097442	Rattus no
	700	AC008723	8.8	16	109616	9	AC008723	Homo sapi
	701	AC114782_3	8.8	16	110000	2	Continuation (4 of	
	702	AL583837_0	8.8	16	110000	8	AL583837	Homo sapi
	703	AL691517_2	8.8	16	110000	9	Continuation (3 of	
	704	LMFLCHR12_1	8.8	16	110000	9	Continuation (2 of	
	705	LMFLCHR12_1	8.8	16	110000	9	LMFLCHR12_1	
	706	LMFLCHR16_06	8.8	16	110000	2	Continuation (7 of	
	707	LMFLCHR16_12	8.8	16	110000	2	Continuation (13 o	
	708	LMFLCHR32_12	8.8	16	111336	2	Continuation (14 o	
	709	LMFLCHR32_13	8.8	16	111464	2	AC018305	Drosophil
c	710	AC018305	8.8	16	112484	8	AC677971	Sequence
c	711	F10K1	8.8	16	112831	2	AC068798	Homo sapi
	712	AC068798	8.8	16	114178	2	AP002336	Homo sapi
	713	AP002336	8.8	16	114502	9	AC004974	Homo sapi
	714	AC004974	8.8	16	115037	2	AL390067	Human DNA
	715	AL390067	8.8	16	115043	9	AC098009	Rattus no
	716	AC098009	8.8	16	115154	2	AC130011	Homo sapi
	717	AC130011	8.8	16	115499	2	AP002822	Homo sapi
	718	AP002822	8.8	16	115562	2	AP003804	Oryza sat
	719	AP003804	8.8	16	116080	9	AC115363	Rattus no
	720	AC115363	8.8	16	116304	2	AC116204	Rattus no
	721	AC116204	8.8	16	116580	8	AC092387	Oryza sat
	722	AC092387	8.8	16	116580	10	AC112162	Mus Muscu

Right group (rows 723–795)

	#			Pos	Cnt	Accession	Species
	723	8.8	16	116620	2	AC123322	Rattus no
	724	8.8	16	116625	2	AC068347	Homo sapi
	725	8.8	16	116691	9	HSJ570O12 / AL121835	Human DNA
c	726	8.8	16	116841	2	AP000643	Homo sapi
	727	8.8	16	117487	2	AP002335	Homo sapi
c	728	8.8	16	118309	2	AP004180	Homo sapi
c	729	8.8	16	118327	8	AC004521	Oryza sat
	730	8.8	16	118684	5	AC004259	Arabidops
	731	8.8	16	119385	5	AL645782	Human Chr
	732	8.8	16	119535	2	AC129382	Zebrafish
c	733	8.8	16	119617	9	AC121015	Rattus no
	734	8.8	16	120662	2	AC105388	Rattus no
	735	8.8	16	120990	2	OSJN01011	Homo sapi
c	736	8.8	16	121080	9	AC120314	Oryza sat
	737	8.8	16	121958	2	AL583838	Rattus no
	738	8.8	16	122012	9	AC008388	Human DNA
	739	8.8	16	122418	5	AL592306	Homo sapi
	740	8.8	16	122427	5	AC025459	Zebrafish
c	741	8.8	16	122437	2	AC022498	Homo sapi
	742	8.8	16	122912	2	AC125783	Homo sapi
c	743	8.8	16	122948	9	HS891H21	Oryza sat
	744	8.8	16	123192	10	AF121351	Human DNA
c	745	8.8	16	123209	8	ATF4F15 / AL049711	Mus muscu / Arabidops
	746	8.8	16	124092	8	CNS08C9K / AL732645	Oryza sat
	747	8.8	16	124377	2	AP005390	Oryza sat
c	748	8.8	16	124861	9	AC010314	Homo sapi
	749	8.8	16	124925	2	AC096184	Homo sapi
c	750	8.8	16	125145	2	AC022136	Homo sapi
c	751	8.8	16	125571	9	AP003703	Oryza sat
c	752	8.8	16	126150	2	AC005152	Oryza sat
	753	8.8	16	126698	2	AC115497	Rattus no
c	754	8.8	16	126756	2	AC124215	Medicago
	755	8.8	16	127334	2	AC109579	Homo sapi
	756	8.8	16	127993	5	AL596026	Zebrafish
	757	8.8	16	128501	2	AC012614	Homo sapi
	758	8.8	16	128763	2	AC126779	Medicago
c	759	8.8	16	128782	9	AC010266	Homo sapi
c	760	8.8	16	128941	9	AC110348	Rattus no
c	761	8.8	16	129547	2	AL589677	Human DNA
c	762	8.8	16	130049	8	AC006434	Genomic s
	763	8.8	16	130244	9	AC073296	Mus muscu
	764	8.8	16	130594	9	AL590705	Human DNA
	765	8.8	16	130730	9	HS63M23 / Z98754	Human DNA s
c	766	8.8	16	131351	2	AC126057	Oryza sat
	767	8.8	16	131481	2	AC016412	Homo sapi
c	768	8.8	16	132623	2	AC119613	Rattus no
	769	8.8	16	133181	9	AC108205	Homo sapi
c	770	8.8	16	133258	9	AC100854	Homo sapi
	771	8.8	16	133847	2	AC004944	Homo sapi
	772	8.8	16	133863	9	AC096666	Homo sapi
	773	8.8	16	133909	9	AC004600	Homo sapi
c	774	8.8	16	133924	2	AC107430	Rattus no
c	775	8.8	16	133925	4	AC093712	Canis fam
	776	8.8	16	134182	2	AC004562	Homo sapi
	777	8.8	16	134350	8	AC067868	Homo sapi
c	778	8.8	16	135062	2	AC117507	Homo sapi
c	779	8.8	16	135689	9	AC105734	Oryza sat
c	780	8.8	16	135809	2	AC026691	Homo sapi
	781	8.8	16	136198	2	AC023755	Homo sapi
	782	8.8	16	136321	9	AC079021	Oryza sat
c	783	8.8	16	136434	2	AL392004	Oryza sat
c	784	8.8	16	136500	9	AP004657	Oryza sat
	785	8.8	16	136776	2	AP003300	Homo sapi
	786	8.8	16	136869	9	AC067868	Homo sapi
	787	8.8	16	137048	2	AC097067	Rattus no
c	788	8.8	16	137122	9	AC127910	Homo sapi
	789	8.8	16	137271	2	AP005193	Mus muscu
c	790	8.8	16	137390	9	AC116724	Rattus no
	791	8.8	16	137716	8	AC115202	Rattus no
	792	8.8	16	137751	10	AC102823	Homo sapi
	793	8.8	16		9	AC112205	Homo sapi
	794	8.8	16		8	AP003310	Oryza sat
	795	8.8	16		10	AC091452	Mus muscu

Left block (rows 796–868):

	#	ID	n		pos		Accession	Species
c	796	AC091452	10	8.8	137751	16	AC091452	Mus muscu
	797	AC095231	2	8.8	138397	16	AC095231	Rattus no
c	798	CNS08C7Y	2	8.8	138477	16	AL731749	Oryza sat
	799	AC114385	2	8.8	138733	16	AC114385	Rattus no
c	800	AC128044	2	8.8	138995	16	AC128044	Rattus no
	801	AC096850	2	8.8	139019	16	AC096850	Pan trogl
	802	AC106510	9	8.8	139054	16	AC106510	Rattus no
c	803	AC097060	2	8.8	139177	16	AC097060	Homo sapi
	804	HSDA60B16	9	8.8	139258	16	AL050305	Human DNA
c	805	AP005592	2	8.8	139616	16	AP005592	Oryza sat
c	806	AC127661	2	8.8	139715	16	AC127661	Rattus no
	807	AP004867	2	8.8	139813	16	AP004867	Oryza sat
	808	H0512B01	8	8.8	140072	16	AL442110	Oryza sat
c	809	AP003276	2	8.8	140491	16	AP003276	Oryza sat
	810	AC018955	9	8.8	141281	16	AC018955	Homo sapi
	811	AP004863	2	8.8	141446	16	AP004863	Oryza sat
	812	AL355353	9	8.8	141605	16	AL355353	Human DNA
	813	AC100353	2	8.8	141714	16	AC100353	Homo sapi
	814	AC124283	2	8.8	141737	16	AC124283	Homo sapi
c	815	AC117174	2	8.8	141798	16	AC117174	Rattus no
	816	AC087421	4	8.8	141881	16	AC087421	Felis cat
	817	AP004705	2	8.8	142024	16	AP004705	Oryza sat
	818	AC098113	2	8.8	142172	16	AC098113	Rattus no
c	819	AC015594	2	8.8	142172	16	AC015594	Homo sapi
c	820	AC023980	2	8.8	142498	16	AC023980	Homo sapi
	821	AP004611	2	8.8	142680	16	AP004611	Oryza sat
	822	AL391599	8	8.8	142762	16	AL391599	Homo sapi
	823	AC098617	2	8.8	142812	16	AC098617	Homo sapi
	824	AC025840	2	8.8	142969	16	AC025840	Homo sapi
	825	AC025840	2	8.8	142969	16	AC025840	Homo sapi
	826	AP001787	2	8.8	143409	16	AP001787	Homo sapi
	827	AP005181	2	8.8	143426	16	AP005181	Oryza sat
	828	AC105391	2	8.8	143498	16	AC105391	Homo sapi
c	829	AC115856	2	8.8	144018	16	AC115856	Mus muscu
	830	AP001766	2	8.8	144041	16	AP001766	Homo sapi
c	831	AC067926	2	8.8	144449	16	AC067926	Mus muscu
c	832	AC025736	2	8.8	144585	16	AC025736	Homo sapi
	833	AF229187	2	8.8	144651	16	AF229187	Oryza sat
	834	AC105768	9	8.8	145003	16	AC105768	Homo sapi
	835	AC021078	9	8.8	145050	16	AC021078	Homo sapi
	836	AL683886	9	8.8	145108	16	AL683886	Human DNA
	837	AC104361	2	8.8	145172	16	AC104361	Homo sapi
	838	AC109644	2	8.8	145267	16	AC109644	Oryza sat
	839	OSJN00243	2	8.8	145320	16	AL731600	Oryza sat
	840	AL135919	2	8.8	145414	16	AL135919	Homo sapi
	841	AC068352	2	8.8	145450	16	AC068352	Homo sapi
	842	AP004270	2	8.8	145467	16	AP004270	Oryza sat
c	843	HS245G19	2	8.8	145538	16	Z92542	Human DNA s
	844	AF030027	14	8.8	145597	16	AF030027	Equine he
	845	AC099499	2	8.8	145787	16	AC099499	Homo sapi
	846	AC004931	2	8.8	145844	16	AC004931	Homo sapi
	847	OSJN00027	2	8.8	145910	16	AL606600	Oryza sat
	848	AP005474	2	8.8	145953	16	AP005474	Oryza sat
	849	AC018833	2	8.8	145976	16	AC018833	Homo sapi
c	850	AC117837	10	8.8	145990	16	AC117837	Mus muscu
c	851	AC116524	2	8.8	146095	16	AC116524	Mus muscu
	852	AC024942	2	8.8	146120	16	AC024942	Homo sapi
	853	AP003479	2	8.8	146340	16	AP003479	Homo sapi
	854	AC083963	2	8.8	146377	16	AC083963	Homo sapi
	855	AC111188	9	8.8	146394	16	AC111188	Homo sapi
	856	AP005258	2	8.8	146425	16	AP005258	Oryza sat
	857	AC128276	2	8.8	146571	16	AC128276	Rattus no
c	858	AC093755	9	8.8	146673	16	AC093755	Homo sapi
	859	AC026931	2	8.8	146916	16	AC026931	Homo sapi
	860	AC026014	2	8.8	147178	16	AC026014	Homo sapi
	861	AC119708	2	8.8	147381	16	AC119708	Rattus no
c	862	AC084020	10	8.8	147670	16	AC084020	Mus muscu
	863	AP001328	2	8.8	147826	16	AP001328	Homo sapi
	864	AC092993	9	8.8	147859	16	AC092993	Homo sapi
	865	AP004234	2	8.8	147887	16	AP004234	Oryza sat
	866	AC026913	2	8.8	147940	16	AC026913	Homo sapi
	867	AB001488	1	8.8	148068	16	AB001488	Bacillus
c	868	AC025476	2	8.8	148069	16	AC025476	Homo sapi

Right block (rows 869–941):

	#	ID	n		pos		Accession	Species
c	869	HS1068E13	9	8.8	148177	16	AL035563	Human DNA
	870	AC109661	2	8.8	148266	16	AC109661	Rattus no
c	871	AL359199	2	8.8	148290	16	AL359199	Human DNA
	872	AC061971	2	8.8	148405	16	AC061971	Homo sapi
	873	AC019072	2	8.8	148438	16	AC019072	Homo sapi
c	874	AC109851	9	8.8	148477	16	AC109851	Rattus no
	875	CNS01RGK	2	8.8	148515	16	AL158059	Human chr
c	876	AC123833	2	8.8	148567	16	AC123833	Mus muscu
	877	AC124213	8	8.8	148704	16	AC124213	Genomic s
	878	AC111845	2	8.8	148883	16	AC111845	Rattus no
c	879	AC128011	2	8.8	148994	16	AC128011	Rattus no
c	880	AP001191	2	8.8	149035	16	AP001191	Homo sapi
	881	AC115273	2	8.8	149405	16	AC115273	Rattus no
c	882	AC027218	9	8.8	149437	16	AC027218	Homo sapi
	883	AC087428	2	8.8	149627	16	AC087428	Homo sapi
	884	AC002076	2	8.8	149764	16	AC002076	Homo sapi
	885	AL773509	9	8.8	150017	16	AL773509	Mus muscu
	886	AL844603	2	8.8	150347	16	AL844603	Mus muscu
c	887	AC027038	2	8.8	150429	16	AC027038	Oryza sat
	888	AC012089	2	8.8	150561	16	AC012089	Homo sapi
	889	AC109543	2	8.8	150561	16	AC109543	Rattus no
	890	AC027101	2	8.8	150561	16	AC027101	Homo sapi
c	891	AC068129	9	8.8	150646	16	AC068129	Homo sapi
c	892	AC016535	2	8.8	150652	16	AC016535	Homo sapi
	893	AC101964	2	8.8	150781	16	AC101964	Mus muscu
	894	AC127815	2	8.8	151042	16	AC127815	Rattus no
c	895	AC012001	2	8.8	151336	16	AC012001	Homo sapi
c	896	AC111505	2	8.8	151446	16	AC111505	Rattus no
	897	AC102022	2	8.8	152346	16	AC102022	Mus muscu
	898	AL138894	2	8.8	152366	16	AL138894	Human DNA
c	899	AC008568	9	8.8	152706	16	AC008568	Homo sapi
c	900	AC008876	9	8.8	152707	16	AC008876	Homo sapi
	901	AC022381	2	8.8	152742	16	AC022381	Homo sapi
	902	AC064852	2	8.8	152781	16	AC064852	Homo sapi
	903	AC079214	9	8.8	153021	16	AC079214	Homo sapi
	904	AC087091	2	8.8	153057	16	AC087091	Homo sapi
	905	AL591855	9	8.8	153214	16	AL591855	Human DNA
	906	AC026206	2	8.8	153258	16	AC026206	Homo sapi
	907	AC025676	2	8.8	153334	16	AC025676	Homo sapi
c	908	AC016316	2	8.8	153586	16	AC016316	Homo sapi
	909	AC013551	2	8.8	153823	16	AC013551	Homo sapi
	910	AC118806	2	8.8	153824	16	AC118806	Rattus no
c	911	AC079294	2	8.8	154007	16	AC079294	Homo sapi
	912	AC095937	2	8.8	154134	16	AC095937	Rattus no
c	913	AC073295	2	8.8	154286	16	AC073295	Mus muscu
c	914	AP005494	2	8.8	154288	16	AP005494	Oryza sat
	915	AC126160	2	8.8	154301	16	AC126160	Rattus no
c	916	AP001767	2	8.8	154313	16	AP001767	Homo sapi
c	917	AC115065	2	8.8	154338	16	AC115065	Mus muscu
	918	AC021279	9	8.8	154413	16	AC021279	Homo sapi
	919	AC090805	2	8.8	154461	16	AC090805	Homo sapi
	920	AL159158	2	8.8	154682	16	AL159158	Human DNA
	921	AL157944	2	8.8	154746	16	AL157944	Human DNA
c	922	AC032035	2	8.8	154840	16	AC032035	Homo sapi
c	923	AL627406	10	8.8	155214	16	AL627406	Mouse DNA
	924	AC123386	2	8.8	155375	16	AC123386	Rattus no
	925	AL831747	2	8.8	155507	16	AL831747	Danio rer
	926	AC120320	2	8.8	155577	16	AC120320	Rattus no
	927	AC012097	2	8.8	155670	16	AC012097	Drosophil
	928	CEY19E4B	2	8.8	155693	16	AL110487	Caenorhab
	929	AC024218	2	8.8	155808	16	AC024218	Homo sapi
	930	AC022806	2	8.8	156194	16	AC022806	Homo sapi
	931	AC093799	2	8.8	156195	16	AC093799	Homo sapi
	932	AC130399	2	8.8	156228	16	AC130399	Homo sapi
	933	AL591720	2	8.8	156491	16	AL591720	Human DNA
	934	AC111000	2	8.8	156578	16	AC111000	Homo sapi
	935	AP004694	2	8.8	156595	16	AP004694	Oryza sat
	936	AP005428	2	8.8	156651	16	AP005428	Oryza sat
c	937	AC114208	2	8.8	156687	16	AC114208	Rattus no
	938	AL359764	2	8.8	156707	16	AL359764	Human DNA
c	939	AC018541	2	8.8	156785	16	AC018541	Homo sapi
c	940	AC007572	3	8.8	156823	16	AC007572	Drosophil
c	941	AC026726	9	8.8	156825	16	AC026726	Homo sapi

c	942	16	8.8	156891	2	AC108550	Rattus no
c	943	16	8.8	157069	9	AC010163	Homo sapi
	945	16	8.8	157191	2	AC116681	Homo sapi
	945	16	8.8	157302	9	AL445209	Human DNA
	946	16	8.8	157361	2	AC106972	Rattus no
	947	16	8.8	157375	9	AC025207	Homo sapi
	948	16	8.8	157383	9	AC021607	Homo sapi
	949	16	8.8	157390	9	AC025366	Homo sapi
c	950	16	8.8	158084	9	AP003492	Oryza sat
	951	16	8.8	158213	2	AF214635	Homo sapi
	952	16	8.8	158271	9	AC097486	Homo sapi
	953	16	8.8	158496	9	AC113170	Homo sapi
c	954	16	8.8	158548	3	PFMAL1P2	Plasmodiu
	955	16	8.8	158576	9	AC099059	Homo sapi
	956	16	8.8	158792	8	AP003833	Oryza sat
	957	16	8.8	159240	2	AC021309	Homo sapi
	958	16	8.8	159601	2	AC123524	Oryza sat
	959	16	8.8	159880	2	AC101809	Mus muscu
	960	16	8.8	159956	9	AC093289	Oryza sat
	961	16	8.8	160097	2	AC119850	Mus muscu
	962	16	8.8	160359	2	AC119584	Rattus no
	963	16	8.8	160363	9	AP003526	Oryza sat
	964	16	8.8	160394	7	AC012273	Homo sapi
	965	16	8.8	160481	2	AC021784	Homo sapi
	966	16	8.8	160569	9	AC015878	Homo sapi
c	967	16	8.8	160624	9	AC060835	Homo sapi
	968	16	8.8	160759	2	AC117082	Dictyoste
	969	16	8.8	160799	2	AC021601	Homo sapi
c	970	16	8.8	160813	9	AC099502	Homo sapi
	971	16	8.8	160863	9	AC123408	Rattus no
	972	16	8.8	161012	2	AP004277	Oryza sat
	973	16	8.8	161043	10	AL589651	Mouse DNA
	974	16	8.8	161079	2	AC083842	Homo sapi
	975	16	8.8	161247	9	AL390247	Human DNA
	976	16	8.8	161298	2	AP002858	Homo sapi
	977	16	8.8	161464	9	AC104096	Mus muscu
	978	16	8.8	161478	9	AC025538	Homo sapi
	979	16	8.8	161509	9	AC008243	Homo sapi
	980	16	8.8	161674	10	AC093445	Mus muscu
	981	16	8.8	161737	2	AC090512	Homo sapi
	982	16	8.8	161752	9	AC111862	Rattus no
	983	16	8.8	161839	2	AC097391	Rattus no
	984	16	8.8	161964	2	AC112751	Rattus no
	985	16	8.8	162010	2	AP004121	Oryza sat
	986	16	8.8	162045	9	AC114643	Mus muscu
	987	16	8.8	162130	2	AL359237	Human chr
	988	16	8.8	162191	9	AC021004	Homo sapi
	989	16	8.8	162337	2	AC109458	Homo sapi
	990	16	8.8	162364	9	AC025354	Homo sapi
	991	16	8.8	162382	2	AP005213	Homo sapi
	992	16	8.8	162433	2	AL773542	Danio rer
	993	16	8.8	162489	2	AC024377	Homo sapi
	994	16	8.8	162514	9	AP005414	Oryza sat
	995	16	8.8	162591	9	AC104460	Homo sapi
	996	16	8.8	162658	9	AC079354	Homo sapi
c	997	16	8.8	162774	2	AL391839	Human DNA
c	998	16	8.8	162864	2	AC034145	Homo sapi
	999	16	8.8			AC122534	Mus muscu
	1000	16	8.8			AC083960	Homo sapi

```
ALIGNMENTS

RESULT 1
AX406939
LOCUS       AX406939        732 bp    DNA    linear   PAT 14-JUN-2002
DEFINITION  Sequence 1 from Patent WO0222875.
ACCESSION   AX406939
VERSION     AX406939.1   GI:21439814
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 1 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
source      1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
CDS         74..445
            /note="mink-related peptide 1, Val substituted for Ala at
            amino acid 116"
            /codon_start=1
            /protein_id="CAD35310.1"
            /db_xref="GI:21439815"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
            NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
            LEESKATIHENIGVAGFKMSP"
variation   420
            /note="The drug associated here was was quinidine."
            /replace="c"
BASE COUNT  221 a   151 c   157 g   203 t
ORIGIN
Query Match            100.0%;  Score 181;  DB 6;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 5e-93;
Matches   181;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCGAAGGATTTTATTACTTAT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCGAAGGATTTTATTACTTAT  139

Qy   61 ATGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  199

Qy  121 GCTGAGAACTTCACTATGTGTACCTCATGCTGATGATTGGAATGTTCTCTTTC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  200 GCTGAGAACTTCACTATGTCATCCTGTACCTCATGTGTGATGATTGGAATGTTCTCTTTC  259

Qy  181 A 181
         |
Db  260 A 260

RESULT 2
AX406941
LOCUS       AX406941        732 bp    DNA    linear   PAT 14-JUN-2002
DEFINITION  Sequence 3 from Patent WO0222875.
ACCESSION   AX406941
VERSION     AX406941.1   GI:21439816
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 3 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
source      1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
CDS         74..445
            /note="mink-related peptide 1, wild type"
            /codon_start=1
            /protein_id="CAD35180.1"
            /db_xref="GI:21439817"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
```

```
                 NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                 LEESKATIHENIGAAGFKMSP"
BASE COUNT     221 a    152 c    157 g    202 t
ORIGIN

Query Match        100.0%;    Score 181;  DB 6;   Length 732;
Best Local Similarity  100.0%;  Pred. No. 5e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy  61 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy 121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTTGGAATGTTCTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTTGGAATGTTCTCTTTC 259

Qy 181 A 181
         |
Db 260 A 260


RESULT 3
AF071002
LOCUS        AF071002       732 bp    mRNA    linear    PRI 29-APR-1999
DEFINITION   Homo sapiens mink-related peptide 1 mRNA, complete cds.
ACCESSION    AF071002
VERSION      AF071002.1  GI:4704422
KEYWORDS     .
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 732)
AUTHORS      Abbott,G.W., Sesti,F., Splawski,I., Buck,M.E., Lehmann,M.H.,
             Timothy,K.W., Keating,M.T. and Goldstein,S.A.
TITLE        MiRP1 forms IKr potassium channels with HERG and is associated with
             cardiac arrhythmia
JOURNAL      Cell 97 (2), 175-187 (1999)
MEDLINE      99235979
PUBMED       10219239
REFERENCE    2  (bases 1 to 732)
AUTHORS      Abbott,G.W., Sesti,F., Buck,M.E. and Goldstein,S.A.N.
TITLE        Direct Submission
JOURNAL      Submitted (05-JUN-1998) Section of Developmental Biology and
             Biophysics, Department of Pediatrics and Boyer Center for Molecular
             Medicine, Yale University School of Medicine, 295 Congress Avenue,
             New Haven, CT 06536, USA
FEATURES     Location/Qualifiers
     source       1..732
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="21"
                  /map="21q11.1"
                  /tissue_type="heart"
     CDS          74..445
                  /note="potassium channel subunit; MiRP1"
                  /codon_start=1
                  /product="mink-related peptide 1"
                  /protein_id="AAD28086.1"
                  /db_xref="GI:4704423"
                  /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                  NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                  LEESKATIHENIGAAGFKMSP"
BASE COUNT     221 a    152 c    157 g    202 t
ORIGIN

Query Match        100.0%;    Score 181;  DB 9;   Length 732;
Best Local Similarity  100.0%;  Pred. No. 5e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy  61 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 199

Qy 121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTTGGAATGTTCTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTTGGAATGTTCTCTTTC 259

Qy 181 A 181
         |
Db 260 A 260


RESULT 4
AF302095
LOCUS        AF302095       809 bp    mRNA    linear    PRI 14-SEP-2000
DEFINITION   Homo sapiens voltage-gated K+ channel subunit MIRP1 (KCNE2) mRNA,
             complete cds.
ACCESSION    AF302095
VERSION      AF302095.1  GI:10121887
KEYWORDS     .
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 809)
AUTHORS      Domenech,A., Estivill,X. and de la Luna,S.
TITLE        Cloning of human MIRP1 cDNA
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 809)
AUTHORS      Domenech,A., Estivill,X. and de la Luna,S.
TITLE        Direct Submission
JOURNAL      Submitted (01-SEP-2000) Medical and Molecular Genetics Center,
             Institut Recerca Oncologica, Avia. de Castelldefels Km 2,7,
             L'Hospitalet de Llobregat, Barcelona 08907, Spain
FEATURES     Location/Qualifiers
     source       1..809
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="21"
                  /map="21q22"
     gene         1..809
                  /gene="KCNE2"
     CDS          141..512
                  /gene="KCNE2"
                  /codon_start=1
                  /product="voltage-gated K+ channel subunit MIRP1"
                  /protein_id="AAG13416.1"
                  /db_xref="GI:10121888"
                  /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                  NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                  LEESKATIHENIGAAGFKMSP"
BASE COUNT     247 a    172 c    189 g    200 t    1 others
ORIGIN

Query Match        100.0%;    Score 181;  DB 9;   Length 809;
Best Local Similarity  100.0%;  Pred. No. 5e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 147 ACTTTATCCAATTCACACAGACGCTGGAAGAGCGTCTTCCGAAGGATTTTTATTACTTAT 206

Qy  61 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 207 ATGGACAATTGGGCGCCAGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 266
```

```
Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 GCTGAGAACTTCTACTATGTCATCCTGTACCTCGTACCTCATGGTGATGATTGGAATGTTCTTTTC 326

Qy    181 A 181
          |
Db    327 A 327
```

RESULT 5
AP000320
LOCUS   AP000320   24608 bp   DNA   linear   PRI 20-NOV-1999
DEFINITION Homo sapiens genomic DNA, chromouome 21q22.1, D21S226-AML region,
            clone:Q12C8, complete sequence.
ACCESSION   AP000320
VERSION   AP000320.1 GI:4835689
KEYWORDS   HTG.
SOURCE   Homo sapiens DNA, clone:Q12C8.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 24608)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE   Homo sapiens 24,608bp genomic DNA of 21q22.1
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2 (bases 1 to 24608)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT   The sequence is a part of the data (ACCESSION No. AP000165 -
            AP000173).
            The sequencing project is supported by Japan Science Technology
            Corporation (JST) and The Institute of Physical and Chemical
            Research (RIKEN).
FEATURES   Location/Qualifiers
     source   1..24608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="Q12C8"
BASE COUNT   6797 a   5420 c   5437 g   6954 t
ORIGIN

Query Match   100.0%; Score 181; DB 9; Length 24608;
Best Local Similarity   100.0%; Pred. No. 4.6e-93;
Matches   181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  15702 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 15761

Qy     61 ATGGACAATTGGCGCGCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  15762 ATGGACAATTGGCGCGCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCAAGTTGAT 15821

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  15822 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTTC 15881

Qy    181 A 181
          |
Db  15882 A 15882
```

RESULT 6

AP000052
LOCUS   AP000052   100000 bp   DNA   linear   PRI 20-NOV-1999
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, segment 23/28,
            complete sequence.
ACCESSION   AP000052
VERSION   AP000052.1 GI:3132362
KEYWORDS   HTG.
SOURCE   Homo sapiens DNA, clone:245P17-f4A4f_2.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-MAY-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   This sequence is conducted by Kitasato University JST sequencing
            Laboratory as a JST sequencing team.
            Principal Investigator:Yoshiyuki Sakaki Ph.D.
            Phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
            sakaki@hgc.ims.u-tokyo.ac.jp
            Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The
            sequence is submitted by:Human Genome Sequencing in ALIS project of
            JST
            Japan Science and Technology Corporation (JST)
            5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www-alis.tokyo.jst.go.jp/HGS/top.html)
            or send email to webmaster@www-alis.tokyo.jst.go.jp'.
FEATURES   Location/Qualifiers
     source   1..100000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="245P17-f4A4f_2"
BASE COUNT   27603 a 21934 c 22513 g 27950 t
ORIGIN

Query Match   100.0%; Score 181; DB 9; Length 100000;
Best Local Similarity   100.0%; Pred. No. 4.5e-93;
Matches   181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80224 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 80283

Qy     61 ATGGACAATTGGCGCGCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80284 ATGGACAATTGGCGCGCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCAAGTTGAT 80343

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  80344 GCTGAGAACTTCTACTATGTCATCCTGTACCTCGTACCTCATGGTGATGATTGGAATGTTCTTTTC 80403

Qy    181 A 181
          |
Db  80404 A 80404
```

RESULT 7
AP000167
LOCUS   AP000167   100000 bp   DNA   linear   PRI 08-JAN-2000
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
            clone B2344F14-f50E8, segment 3/9, complete sequence.
ACCESSION   AP000167

```
VERSION     AP000167.1  GI:4827132
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Homo sapiens 890,291bp genomic DNA of 21q22.1 (REGION: D21S226-AML
            CLONE RANGE: B2344F14-f50E8)
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     E. coli transposon insertion:The present data does not contain E.
            coli transposon sequences which integrated in the
            original/previous sequences. We determined the boundary between
            the insertion and genomic sequences experimentally, removed the
            insertion sequences, reconstituted the present data. The sequencing
            project is supported by Japan Science Technology Corporation (JST)
            and The Institute of Physical and Chemical Research (RIKEN).
FEATURES            Location/Qualifiers
     source          1..100000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
BASE COUNT    27338 a  22306 c  22571 g  27785 t
ORIGIN

Query Match           100.0%;  Score 181;  DB 9;  Length 100000;
Best Local Similarity 100.0%;  Pred. No. 4.5e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ACTTTATCCAATTTCACACAGACGCTGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72192 ACTTTATCCAATTTCACACAGACGCTGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 72251

Qy     61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72252 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 72311

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  72312 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 72371

Qy    181 A 181
          |
Db  72372 A 72372

RESULT 8
AP000120
ID  AP000120   standard; DNA; HUM; 100000 BP.
XX
AC  AP000120;
XX
SV  AP000120.1
XX
DT  04-MAY-1999 (Rel. 59, Created)
DT  26-SEP-1999 (Rel. 61, Last updated, Version 3)
XX
DE  Homo sapiens genomic DNA of 21q22.1, GART and AML related, SLC5A3-f4A4
DE  region, segment 3/8.
XX
```

```
KW   HTG.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-100000
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RT   ;
RL   Submitted (15-APR-1999) to the EMBL/GenBank/DDBJ databases.
RL   Mika Hirakawa, Japan Science and Technology Corporation (JST), Advanced
RL   Databases Department; 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081, Japan
RL   (E-mail:mika@tokyo.jst.go.jp, URL:http://www-alis.tokyo.jst.go.jp/,
RL   Tel:81-3-5214-8491, Fax:81-3-5214-8470)
XX
RN   [2]
RA   Hattori M., Ishii K., Toyoda A., Shiba T., Sakaki Y.;
RT   "Homo sapiens 817,199bp genomic DNA of 21q22.1 GART and AML region";
RL   Unpublished.
XX
CC   This sequence is conducted by Kitasato University JST sequencing
CC   Laboratory as a JST sequencing team.
CC   Principal Investigator:Yoshiyuki Sakaki Ph.D.
CC   Phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
CC   sakaki@hgc.ims.u-tokyo.ac.jp
CC   Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D.
CC   The sequence is submitted by Human Genome Sequencing in ALIS
CC   project of JST.
CC   Japan Science and Technology Corporation (JST)
CC   5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081 Japan
CC   For further information about this sequence, including its
CC   location and relationship to other sequences, please visit our
CC   sequence archive Web site (http://www-alis.tokyo.jst.go.jp/HGS/)
CC   or send email to webmaster@www-alis.tokyo.jst.go.jp
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..100000
FT                   /chromosome="21"
FT                   /db_xref="taxon:9606"
FT                   /sequenced_mol="DNA"
FT                   /organism="Homo sapiens"
FT                   /map="21q22.1"
FT   STS             3213..32320
FT                   /note="83F9aL-F/83F9aL-R:The location is between each
FT                   flanking site of PCR primers."
FT                   /standard_name="D21S1689"
FT   STS             43934..44157
FT                   /note="AFMa086yf9;Genethon Marker;The location is between
FT                   each flanking site of PCR primers."
FT                   /standard_name="D21S1920"
XX
SQ   Sequence 100000 BP; 27731 A; 21990 C; 22220 G; 28059 T; 0 other;

Query Match           100.0%;  Score 181;  DB 17;  Length 100000;
Best Local Similarity 100.0%;  Pred. No. 4.5e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ACTTTATCCAATTTCACACAGACGCTGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89526 ACTTTATCCAATTTCACACAGACGCTGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 89585

Qy     61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89586 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAAGTTGAT 89645

Qy    121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89646 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTC 89705

Qy    181 A 181
          |
```

```
Db      89706 A 89706

RESULT 9
AP001719
LOCUS       AP001719              340000 bp    DNA     linear   PRI 10-MAY-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 21q, section 63/105.
ACCESSION   AP001719 AL163264 BA000005
VERSION     AP001719.1  GI:7768719
KEYWORDS    .
SOURCE      Homo sapiens DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
AUTHORS     Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., and
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
TITLE       The DNA sequence of human chromosome 21
JOURNAL     Nature 405 (6784), 311-319 (2000)
MEDLINE     20289799
REFERENCE   2  (bases 1 to 340000)
AUTHORS     Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., and
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
TITLE       Direct Submission
JOURNAL     Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing
            Consortium.* RIKEN Genomic Sciences Center, Human Genome Research
            Group * Institute of Molecular Biotechnology, Genome Analysis *
            Keio University School of Medicine, Dept. of Molecular Biology *
            GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
            Genetics (addresses see below)
COMMENT     On May 30, 2000 this sequence version replaced gi:7717333.
            The chromosome 21 mapping and sequencing consortium consisting of
            * RIKEN Genomic Sciences Center, Human Genome Research Group, *
            Sagamihara 228-8555, Japan.
            * e.mail: hattori@gsc.riken.go.jp
            * URL: http://hgp.gsc.riken.go.jp/
            and
            * Institute of Molecular Biotechnology, Genome Analysis, *
            Beutenbergstrasse 11, D-07745 Jena, Germany,
            * e.mail: genome@genome.imb-jena.de
            * URL: http://genome.imb-jena.de/
            and
            * Keio University School of Medicine, Molecular Biology, * Tokyo
            160-8582, Japan,
            * e.mail: nshimizu@dmb-med.keio.ac.jp
            * URL: http://www.dmb.med.keio.ac.jp/
            and
            * GBF, Dept. of Genome Analysis,
            * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
            info.genome@gbf.de
            * URL: http://genome.gbf.de/
            and
            * Max-Planck Institute for Molecular Genetics,
            * Ihnestrasse 73, D-14195 Berlin, Germany,
            * e.mail: info-chr21@molgen.mpg.de
            * URL: http://chr21.rz-berlin.mpg.de/
            AL163264: Submitted (10-Apr-2000).
FEATURES             Location/Qualifiers
     source          1..340000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     source          <1..29420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R338L7, 5' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000467"
     source          29242..97750
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="CTD344F14"
                     /clone_lib="CTD BAC library"
                     /note="Accession No. AP000315"
     source          96883..121400
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pS185"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000316"
     source          121210..218631
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="P245P17"
                     /clone_lib="RP1 PAC library"
                     /note="Accession No. AP000317"
     source          218558..260079
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ6f8"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000318"
     source          258735..285904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ94A10"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000319"
     source          285732..310339
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ12C8"
                     /clone_lib="CMP21 P1 library"
                     /note="Accession No. AP000320"
     source          310050..326194
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="pQ82F5"
```

```
                /clone_lib="CMP21 P1 library"
                /note="Accession No. AP000321"
                325966..>340000
source          /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="21"
                /map="21q22.1"
                /clone="pQ9708, 3' partial"
                /clone_lib="CMP21 P1 library"
                /note="Accession No. AP000322"
repeat_region   complement(1182..1688)
                /note="AluSp"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(2099..2390)
                /note="AluSx"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   2799..2884
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(2891..3287)
                /note="MER31A"
                /rpt_family="LTR/MER4-group"
                /rpt_type=DISPERSED
repeat_region   3288..3381
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
misc_feature    4211..5231
                /note="CpG island"
misc_feature    4272..4279
                /note="NotI site"
repeat_region   4328..4476
                /note="(CGGG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
gene            4392..35127
                /gene="SLC5A3"
mRNA            join(4392..4740,12254..35127)
                /gene="SLC5A3"
                /product="Sodium/myo-inositol cotransporter"
                /note="human solute carrier family 5, member 3, Accession
                No. AF027153"
exon            4392..4740
                /gene="SLC5A3"
                /number=1
misc_feature    4392..4399
                /gene="SLC5A3"
                /note="NotI site"
repeat_region   4810..4985
                /note="(CGGG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   5099..5191
                /note="(CCG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
```

```
Query Match          100.0%;  Score 181;  DB 9;  Length 340000;
Best Local Similarity 100.0%;  Pred. No. 4.4e-93;
Matches 181;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301433 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 301492

Qy      61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301493 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGAT 301552

Qy     121 GCTGAGAACTTCTACTATGTCATCCTGCTACCTCATGGTGATGATTGGAATGTTCTCTTTC  180
```

```
RESULT 10
AX406947
LOCUS        AX406947              732 bp    DNA      linear    PAT 14-JUN-2002
DEFINITION   Sequence 9 from Patent WO022875.
ACCESSION    AX406947
VERSION      AX406947.1    GI:21439822
KEYWORDS     .
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Goldstein,S.A.
  TITLE      Polymorphisms associated with cardiac arrythmia
  JOURNAL    Patent: WO 022875-A 9 21-MAR-2002;
             YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="mink-related peptide 1, Ala substituted for Thr at
                     amino acid 8"
                     /codon_start=1
                     /protein_id="CAD35312.1"
                     /db_xref="GI:21439823"
                     /translation="MSTLSNFAQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
     variation       95
                     /note="The drug associated here is was Bactrim."
                     /replace="a"

BASE COUNT     220 a    152 c    158 g    202 t
ORIGIN
```

```
Query Match          91.2%;  Score 165;  DB 6;  Length 732;
Best Local Similarity 100.0%;  Pred. No. 8.4e-84;
Matches 165;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      17 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTATTATATGGACAATTGGCGCC   76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96 CACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTATTATATGGACAATTGGCGCC  155

Qy      77 AGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT  136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156 AGAACACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACT  215

Qy     137 ATGTCATCCTGTACCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA  181
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     216 ATGTCATCCTGTACCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA  260
```

```
RESULT 11
AX406945
LOCUS        AX406945              732 bp    DNA      linear    PAT 14-JUN-2002
DEFINITION   Sequence 7 from Patent WO022875.
ACCESSION    AX406945
VERSION      AX406945.1    GI:21439820
KEYWORDS     .
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Goldstein,S.A.
```

Left column:

```
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 7 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
  source    1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
  CDS       74..445
            /note="minK-related peptide 1, Thr substituted for Ile at
            amino acid 57"
            /codon_start=1
            /protein_id="CAD35311.1"
            /db_xref="GI:21439821"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
            NFYYVILYLTVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
            LEESKATIHENIGAAGFKMSP"
  variation 243
            /note="The drug associated here was oxatomide."
            /replace="t"
BASE COUNT  221 a       153 c       157 g       201 t
ORIGIN
```

```
Query Match            90.1%;     Score 163;  DB 6;  Length 732;
Best Local Similarity  100.0%;    Pred. No. 1.2e-82;
Matches  163;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 199

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 163
        ||||||||||||||||||||||||||||||||||||||||||
Db  200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGA 242
```

```
RESULT 12
AX406943
LOCUS       AX406943                  732 bp    DNA       linear    PAT 14-JUN-2002
DEFINITION  Sequence 5 from Patent WO0222875.
ACCESSION   AX406943
VERSION     AX406943.1   GI:21439818
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 5 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES    Location/Qualifiers
  source    1..732
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
  CDS       74..445
            /note="minK-related peptide 1, Thr substitued for Met at
            amino acid 54"
            /codon_start=1
            /protein_id="CAD35181.1"
            /db_xref="GI:21439819"
            /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
            NFYYVILYLTVMIGMFSFIIVAILVSTVKSKREHSNDPYHQYIVEDWQEKYKSQILN
            LEESKATIHENIGAAGFKMSP"
  variation 234
            /note="The drug associated here was procainamide."
            /replace="t"
BASE COUNT  221 a       153 c       157 g       201 t
ORIGIN
```

Right column:

```
Query Match            85.1%;     Score 154;  DB 6;  Length 732;
Best Local Similarity  100.0%;    Pred. No. 1.8e-77;
Matches  154;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTTATTACTTAT 139

Qy   61 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 ATGGACAATTGGCGCCAGAACAACAACAGCTGAGCAAGAGAGGCCCTCCAAGCCAAGTTGAT 199

Qy  121 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 154
        |||||||||||||||||||||||||||||||||
Db  200 GCTGAGAACTTCTACTATGTCATCCTGTACCTCA 233
```

```
RESULT 13
AY050513
LOCUS       AY050513                  372 bp    mRNA      linear    ROD 15-OCT-2001
DEFINITION  Cavia porcellus minK-related peptide 1 mRNA, complete cds.
ACCESSION   AY050513
VERSION     AY050513.1   GI:16151156
KEYWORDS    .
SOURCE      Cavia porcellus.
ORGANISM    Cavia porcellus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
REFERENCE   1  (bases 1 to 372)
AUTHORS     Jiang,M., Zhang,M., Liu,J. and Tseng,G.-N.
TITLE       Direct Submission
JOURNAL     Submitted (12-AUG-2001) Physiology, Virginia Commonwealth
            University, 1101 East Marshall Street, Richmond, VA 23298, USA
FEATURES    Location/Qualifiers
  source    1..372
            /organism="Cavia porcellus"
            /db_xref="taxon:10141"
            /tissue_type="heart"
  CDS       1..372
            /note="Mirp1"
            /codon_start=1
            /product="minK-related peptide 1"
            /protein_id="AAL13163.1"
            /db_xref="GI:16151157"
            /translation="MTTLANLTQTLEDAFKKVFITYMDSWRRNTTAEQQALQARVDAE
            NFYYVILYLMVMIGMFAFIVVAILVSTVKSKREHSQDPYHQYIVEDWQQKYRSQILH
            LEDSKATIHENLGATGFTVSP"
BASE COUNT   95 a        98 c       105 g        74 t
ORIGIN
```

```
Query Match            14.4%;     Score 26;   DB 10;  Length 372;
Best Local Similarity  100.0%;    Pred. No. 0.0012;
Matches  26;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  139 GTCATCCTGTACCTCATGGTGATGAT 164
        ||||||||||||||||||||||||||
Db  145 GTCATCCTGTACCTCATGGTGATGAT 170
```

```
RESULT 14
AF071003
LOCUS       AF071003                  468 bp    mRNA      linear    ROD 29-APR-1999
DEFINITION  Rattus norvegicus minK-related peptide 1 mRNA, complete cds.
ACCESSION   AF071003
VERSION     AF071003.1   GI:4704424
KEYWORDS    .
SOURCE      Rattus norvegicus.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 468)
```

```
AUTHORS     Abbott,G.W., Sesti,F., Splawski,I., Buck,M.E., Lehmann,M.H.,
            Timothy,K.W., Keating,M.T. and Goldstein,S.A.
TITLE       MiRP1 forms IKr potassium channels with HERG and is associated with
            cardiac arrhythmia
JOURNAL     Cell 97 (2), 175-187 (1999)
MEDLINE     9935979
PUBMED      10219239
REFERENCE   2  (bases 1 to 468)
AUTHORS     Abbott,G.W., Sesti,F., Buck,M.E. and Goldstein,S.A.N.
TITLE       Direct Submission
JOURNAL     Submitted (05-JUN-1998) Section of Developmental Biology and
            Biophysics, Department of Pediatrics and Boyer Center for Molecular
            Medicine, Yale University School of Medicine, 295 Congress Avenue,
            New Haven, CT 06536, USA
FEATURES             Location/Qualifiers
     source          1..468
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /tissue_type="heart"
                     /dev_stage="10-12 weeks"
     CDS             35..406
                     /note="potassium channel subunit; MiRP1"
                     /codon_start=1
                     /product="minK-related peptide 1"
                     /protein_id="AAD28087.1"
                     /db_xref="GI:4704425"
                     /translation="MTTLANLTQTLEDAFKKVFITYMDSWRRNTTAEQQAIQARVDAE
                     NFYYVILYLMVMIGMFAFIVVAILVSTVKSKREHSQDPYHQYIVEDWQQKYRSQIIH
                     LEDSKATIHENLGATGFTVSP"
BASE COUNT       118 a       131 g        93 t
ORIGIN          126 c
```

```
Query Match             14.4%;    Score 26;  DB 10;   Length 468;
Best Local Similarity  100.0%;    Pred. No. 0.0012;
Matches  26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps    0;

Qy     139 GTCATCCTGTACCTCATGGTGATGAT 164
           ||||||||||||||||||||||||||
Db     179 GTCATCCTGTACCTCATGGTGATGAT 204
```

```
RESULT 15
BC022699
LOCUS       BC022699         1664 bp    mRNA      linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, RIKEN cDNA 2200002I16 gene, clone MGC:31447
            IMAGE:4481325, mRNA, complete cds.
ACCESSION   BC022699
VERSION     BC022699.1  GI:18490550
KEYWORDS    MGC.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1664)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (01-FEB-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site:      http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriquez, A., and Myers,
```

R. M.

```
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 44 Row: a Column: 22
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1..1664
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="MGC:31447 IMAGE:4481325"
                     /tissue_type="Eye, retina, mouse strain C57Bl\6"
                     /clone_lib="NIH_MGC_94"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             90..461
                     /codon_start=1
                     /product="RIKEN cDNA 2200002I16 gene"
                     /protein_id="AAH22699.1"
                     /db_xref="GI:18490551"
                     /db_xref="LocusID:69143"
                     /translation="MATLANLTQTLEDAFKKVFITYMDSWRRNTTAEEQALQARVDAE
                     NFYYVILYLMVMIGMFSFIVVAILVSTVKSKREHSQHPYHQYIVEDWQEKYKSQIIH
                     LEDSKATIHENMGATGFTVSP"
BASE COUNT       502 a       337 c       404 g       421 t
ORIGIN
```

```
Query Match             14.4%;    Score 26;  DB 10;   Length 1664;
Best Local Similarity  100.0%;    Pred. No. 0.0012;
Matches  26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps    0;

Qy     139 GTCATCCTGTACCTCATGGTGATGAT 164
           ||||||||||||||||||||||||||
Db     234 GTCATCCTGTACCTCATGGTGATGAT 259
```

```
RESULT 16
AC117904/c
LOCUS       AC117904             144709 bp    DNA         linear   HTG 18-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-409A1, *** SEQUENCING IN PROGRESS
            ***, 44 unordered pieces.
ACCESSION   AC117904
VERSION     AC117904.3  GI:21746840
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 144709)
AUTHORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,J., Hollins,B.,
            Homsi,F., Howard,J., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,U., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,J.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
```

```
          Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
          Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
          Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
          Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
          Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
          Moser,M., Neal,D., Newtson,J., Nawtson,N., Nguyen,A., Nguyen,N.,
          Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
          Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
          Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
          Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
          Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
          Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
          Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
          Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,D.,
          Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
          Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
          Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 144709)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 144709)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 14, 2002 this sequence version replaced gi:20162905.
            ----------------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ----------------------- Project Information
            Center project name: GVCH
            Center clone name: CH230-409A1
            ----------------------- Summary Statistics
            Sequencing vector: plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 112047 bases at least Q40
            Consensus quality: 117676 bases at least Q30
            Consensus quality: 121024 bases at least Q20
            ----------------------- Genome Center
          * NOTE: Estimated insert size may differ from sequence length
          * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 44 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *      1    1015: contig of 1015 bp in length
          *   1016    1115: gap of unknown length
          *   1116    2338: contig of 1223 bp in length
          *   2339    2438: gap of unknown length
          *   2439    3686: contig of 1248 bp in length
          *   3687    3786: gap of unknown length
          *   3787    5609: contig of 1823 bp in length
          *   5610    5709: gap of unknown length
          *   5710    7138: contig of 1429 bp in length
          *   7139    7238: gap of unknown length
          *   7239    9288: contig of 2050 bp in length
          *   9289    9388: gap of unknown length
          *   9389   11561: contig of 2173 bp in length
          *  11562   11661: gap of unknown length
          *  11662   12758: contig of 1097 bp in length
          *  12759   12858: gap of unknown length
          *  12859   15068: contig of 2210 bp in length
          *  15069   15168: gap of unknown length
          *  15169   16496: contig of 1328 bp in length
          *  16497   16596: gap of unknown length
          *  16597   19364: contig of 2768 bp in length
          *  19365   19464: gap of unknown length
          *  19465   20896: contig of 1432 bp in length
          *  20897   20996: gap of unknown length
          *  20997   22441: contig of 1445 bp in length
          *  22442   22541: gap of unknown length
          *  22542   24534: contig of 1993 bp in length
          *  24535   24634: gap of unknown length
          *  24635   27140: contig of 2506 bp in length
          *  27141   27240: gap of unknown length
          *  27241   29139: contig of 1899 bp in length
          *  29140   29239: gap of unknown length
          *  29240   31805: contig of 2566 bp in length
          *  31806   31905: gap of unknown length
          *  31906   34154: contig of 2249 bp in length
          *  34155   34254: gap of unknown length
          *  34255   36644: contig of 2390 bp in length
          *  36645   36744: gap of unknown length
          *  36745   37831: contig of 1087 bp in length
          *  37832   37931: gap of unknown length
          *  37932   39841: contig of 1910 bp in length
          *  39842   39941: gap of unknown length
          *  39942   42387: contig of 2446 bp in length
          *  42388   42487: gap of unknown length
          *  42488   45137: contig of 2650 bp in length
          *  45138   45237: gap of unknown length
          *  45238   48966: contig of 3729 bp in length
          *  48967   49066: gap of unknown length
          *  49067   51971: contig of 2905 bp in length
          *  51972   52071: gap of unknown length
          *  52072   54562: contig of 2491 bp in length
          *  54563   54662: gap of unknown length
          *  54663   56481: contig of 1819 bp in length
          *  56482   56581: gap of unknown length
          *  56582   59476: contig of 2895 bp in length
          *  59477   59576: gap of unknown length
          *  59577   63124: contig of 3548 bp in length
          *  63125   63224: gap of unknown length
          *  63225   67624: contig of 4400 bp in length
          *  67625   67724: gap of unknown length
          *  67725   71583: contig of 3859 bp in length
          *  71584   71683: gap of unknown length
          *  71684   75653: contig of 3970 bp in length
          *  75654   75753: gap of unknown length
          *  75754   79243: contig of 3490 bp in length
          *  79244   79343: gap of unknown length
          *  79344   82996: contig of 3653 bp in length
          *  82997   83096: gap of unknown length
          *  83097   88441: contig of 5345 bp in length
          *  88442   88541: gap of unknown length
          *  88542   92876: contig of 4335 bp in length
          *  92877   92976: gap of unknown length
          *  92977   97178: contig of 4202 bp in length
          *  97179   97278: gap of unknown length
          *  97279  103933: contig of 6655 bp in length
          * 103934  104033: gap of unknown length
          * 104034  109263: contig of 5230 bp in length
          * 109264  109363: gap of unknown length
          * 109364  115319: contig of 5956 bp in length
          * 115320  115419: gap of unknown length
          * 115420  122481: contig of 7062 bp in length
          * 122482  122581: gap of unknown length
          * 122582  132083: contig of 9502 bp in length
          * 132084  132183: gap of unknown length
          * 132184  139601: contig of 7418 bp in length
          * 139602  139701: gap of unknown length
```

```
FEATURES      *    139702    144709: contig of 5008 bp in length.
source                       Location/Qualifiers
                             1..144709
                             /organism="Rattus norvegicus"
                             /db_xref="taxon:10116"
                             /clone="CH230-409A1"
BASE COUNT    36065 a   34426 c   33378 g   36487 t    4353 others
ORIGIN

Query Match          14.4%;    Score 26;   DB 2;   Length 144709;
Best Local Similarity  100.0%;  Pred. No. 0.0011;
Matches     26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     139 GTCATCCTGTACCTCATGGTGATGAT 164
           ||||||||||||||||||||||||||
Db  131572 GTCATCCTGTACCTCATGGTGATGAT 131547

RESULT 17
AF329636
LOCUS       AF329636      215 bp    mRNA    linear    MAM 28-SEP-2001
DEFINITION  Oryctolagus cuniculus K+/pacemaker channel beta subunit mirp1
            (kcne2) mRNA, partial cds.
ACCESSION   AF329636
VERSION     AF329636.1  GI:13194729
KEYWORDS    .
SOURCE      Oryctolagus cuniculus.
ORGANISM    Oryctolagus cuniculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE   1  (bases 1 to 215)
AUTHORS     Yu,H., Wu,J., Potapova,I., Wymore,R.T., Holmes,B., Zuckerman,J.,
            Pan,Z., Wang,H., Shi,W., Robinson,R.B., El-Maghrabi,M.R.,
            Benjamin,W., Dixon,J., McKinnon,D., Cohen,I.S. and Wymore,R.S.
TITLE       MinK-related peptide 1: A beta subunit for the HCN ion channel
            subunit family enhances expression and speeds activation
JOURNAL     Circ. Res. 88 (12), E84-E87 (2001)
MEDLINE     21313430
PUBMED      11420311
REFERENCE   2  (bases 1 to 215)
AUTHORS     Wymore,R.T., Holmes,B.A., Wymore,R.S., Yu,H., Wu,J., Potapova,I.,
            Zuckerman,J., Pan,Z., Wang,H., Shi,W., Robinson,R., El-Maghrabi,R.,
            Benjamin,W., Dixon,J.E., McKinnon,D. and Cohen,I.S.
TITLE       Direct Submission
JOURNAL     Submitted (15-DEC-2000) Biology, The University of Tulsa, 600 S.
            College Av., Tulsa, OK 74104-3189, USA
FEATURES    Location/Qualifiers
source      1..215
            /organism="Oryctolagus cuniculus"
            /db_xref="taxon:9986"
gene        <1..>215
            /gene="kcne2"
CDS         <1..>215
            /gene="kcne2"
            /codon_start=3
            /product="K+/pacemaker channel beta subunit mirp1"
            /protein_id="AAK15527.1"
            /db_xref="GI:13194730"
            /translation="AENFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQ
            YIVEDWQEKYKSQILHFEEAKATIHEN"
BASE COUNT   60 a   58 c   51 g   46 t
ORIGIN

Query Match          12.7%;    Score 23;   DB 4;   Length 215;
Best Local Similarity  100.0%;  Pred. No. 0.066;
Matches     23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     124 GAGAACTTCTACTATGTCATCCT 146
           |||||||||||||||||||||||
Db       6 GAGAACTTCTACTATGTCATCCT 28

RESULT 18
```

```
AY079211
LOCUS       AY079211      225 bp    mRNA    linear    MAM 20-JUL-2002
DEFINITION  Sus scrofa KCNE2 mRNA, partial cds.
ACCESSION   AY079211
VERSION     AY079211.1  GI:21913149
KEYWORDS    .
SOURCE      pig.
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 225)
AUTHORS     Li,Y., Mitchell,K.E., Wymore,R.S. and Freeman,L.C.
TITLE       Granulosa cells express multiple KCNQ and KCNE channel subunits
            (Abstract 1220)
JOURNAL     Biophys. J. 82, 252a (2002)
REFERENCE   2  (bases 1 to 225)
AUTHORS     Li,Y., Wymore,T., Mitchell,K.E., Wymore,R.S. and Freeman,L.C.
TITLE       Direct Submission
JOURNAL     Submitted (19-FEB-2002) Anatomy and Physiology, College of
            Veterinary Medicine, Kansas State University, 228 Coles Hall, 1600
            Denison Ave., Manhattan, KS 66506-5802, USA
FEATURES    Location/Qualifiers
source      1..225
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
CDS         <1..>225
            /codon_start=1
            /product="KCNE2"
            /protein_id="AAM77423.1"
            /db_xref="GI:21913150"
            /translation="DAENFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYH
            QYIVEDWQGKYKSQILNLEESRATIHENLGA"
BASE COUNT   66 a   60 c   54 g   45 t
ORIGIN

Query Match          12.7%;    Score 23;   DB 4;   Length 225;
Best Local Similarity  100.0%;  Pred. No. 0.066;
Matches     23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     154 ATGGTGATGATTGGAATGTTCTC 176
           |||||||||||||||||||||||
Db      37 ATGGTGATGATTGGAATGTTCTC 59

RESULT 19
AF387764
LOCUS       AF387764      228 bp    mRNA    linear    MAM 08-NOV-2001
DEFINITION  Equus caballus minK related peptide 1-like mRNA, partial sequence.
ACCESSION   AF387764
VERSION     AF387764.1  GI:16797919
KEYWORDS    .
SOURCE      Equus caballus.
ORGANISM    Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 228)
AUTHORS     Finley,M.R., Mitchell,K.E., Mitchell,B.M. and Freeman,L.C.
TITLE       Post-translational modification of ERG in horse heart
JOURNAL     FASEB J. 15, A1135 (2001)
REFERENCE   2  (bases 1 to 228)
AUTHORS     Finley,M.R., Li,Y. and Freeman,L.C.
TITLE       Direct Submission
JOURNAL     Submitted (31-MAY-2001) Department of Anatomy and Physiology,
            Kansas State University, 1600 Denison Ave, Manhattan, KS 66506, USA
COMMENT     NCBI staff are still waiting for submitters to provide appropriate
            coding sequence information.
FEATURES    Location/Qualifiers
source      1..228
            /organism="Equus caballus"
            /db_xref="taxon:9796"
            /tissue_type="atrium"
misc_feature  44..>228
            /note="similar to minK related peptide 1; potassium
```

```
BASE COUNT              channel accessory protein; MiRP1; KCNE2"
ORIGIN          68 a     57 c     51 g     51 t      1 others

Query Match             12.7%;  Score 23;  DB 4;  Length 228;
Best Local Similarity  100.0%;  Pred. No. 0.066;
Matches   23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     159 GATGATTGGAATGTTCTCTTCA 181
           |||||||||||||||||||||
Db      43 GATGATTGGAATGTTCTCTTCA 65
```

```
RESULT 20
AX406951
LOCUS        AX406951                 21 bp    DNA        linear   PAT 14-JUN-2002
DEFINITION   Sequence 13 from Patent WO0222875.
ACCESSION    AX406951
VERSION      AX406951.1  GI:21439826
KEYWORDS     .
SOURCE       synthetic construct.
ORGANISM     synthetic construct
             artificial sequences.
REFERENCE    1
AUTHORS      Goldstein,S.A.
TITLE        Polymorphisms associated with cardiac arrythmia
JOURNAL      Patent: WO 0222875-A 13 21-MAR-2002;
             YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
source               1..21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR amplification primer"
BASE COUNT    4 a     7 c     2 g     8 t
ORIGIN
```

```
Query Match             11.6%;  Score 21;  DB 6;  Length 21;
Best Local Similarity  100.0%;  Pred. No. 0.99;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     133 TACTATGTCATCCTGTACCTC 153
           |||||||||||||||||||||
Db       1 TACTATGTCATCCTGTACCTC 21
```

```
RESULT 21
AX406950/c
LOCUS        AX406950                 20 bp    DNA        linear   PAT 14-JUN-2002
DEFINITION   Sequence 12 from Patent WO0222875.
ACCESSION    AX406950
VERSION      AX406950.1  GI:21439825
KEYWORDS     .
SOURCE       synthetic construct.
ORGANISM     synthetic construct
             artificial sequences.
REFERENCE    1
AUTHORS      Goldstein,S.A.
TITLE        Polymorphisms associated with cardiac arrythmia
JOURNAL      Patent: WO 0222875-A 12 21-MAR-2002;
             YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
source               1..20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR amplification primer"
BASE COUNT   10 a     4 c     2 g     4 t
ORIGIN
```

```
Query Match             11.0%;  Score 20;  DB 6;  Length 20;
Best Local Similarity  100.0%;  Pred. No. 3.7;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     160 ATGATTGGAATGTTCTCTTT 179
```

```
Db      20 ATGATTGGAATGTTCTCTTT 1
           ||||||||||||||||||||

RESULT 22
AC108241
LOCUS        AC108241               82479 bp    DNA        linear   HTG 13-JUL-2002
DEFINITION   Rattus norvegicus clone CH230-159K5, *** SEQUENCING IN PROGRESS
             ***, 50 unordered pieces.
ACCESSION    AC108241
VERSION      AC108241.3  GI:21731242
KEYWORDS     HTG; HTGS_PHASE1.
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 82479)
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
             Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
             Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.P.,
             Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
             Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
             Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
             Chen,G., Chen,R., Chen,Z., Chowdhry,I., Chrisopoulos,C.,
             Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
             Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
             Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
             Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
             Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
             Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
             Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
             Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
             Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
             Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
             Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
             Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
             Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
             Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
             Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louligeged,H.,
             Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
             Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
             Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
             Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
             Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
             Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
             Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
             Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
             Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
             Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
             Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
             Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,K., Tang,H.,
             Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
             Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
             Wang,S., Ward-Moore,S., Warren,V., Washington,C., Watlington,S.,
             Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
             Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
             Weinstock,G. and Gibbs,R.
TITLE        Direct Submission
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 82479)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (27-JAN-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
REFERENCE    3  (bases 1 to 82479)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
COMMENT      On Jul 11, 2002 this sequence version replaced gi:18846240.
```

```
--------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
--------- Project Information
    Center project name: GJPX
    Center clone name: CH230-159K5
--------- Summary Statistics
    Sequencing vector: Plasmid;
    Chemistry: Dye-terminator Big Dye: 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 17922 bases at least Q40
    Consensus quality: 19302 bases at least Q30
    Consensus quality: 20254 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
*    (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
*    consists of 50 contigs. The true order of the pieces
*    is not known and their order in this sequence record is
*    arbitrary. Gaps between the contigs are represented as
*    runs of N, but the exact sizes of the gaps are unknown.
*    This record will be updated with the finished sequence
*    as soon as it is available and the accession number will
*    be preserved.

*       1  1615: contig of 1615 bp in length
*    1616  1715: gap of unknown length
*    1716  2945: contig of 1230 bp in length
*    2946  3045: gap of unknown length
*    3046  4078: contig of 1033 bp in length
*    4079  4178: gap of unknown length
*    4179  5526: contig of 1348 bp in length
*    5527  5626: gap of unknown length
*    5627  6717: contig of 1091 bp in length
*    6718  6817: gap of unknown length
*    6818  7843: contig of 1026 bp in length
*    7844  7943: gap of unknown length
*    7944  9033: contig of 1090 bp in length
*    9034  9133: gap of unknown length
*    9134 10725: contig of 1592 bp in length
*   10726 10825: gap of unknown length
*   10826 12206: contig of 1381 bp in length
*   12207 12306: gap of unknown length
*   12307 13744: contig of 1438 bp in length
*   13745 13844: gap of unknown length
*   13845 15153: contig of 1309 bp in length
*   15154 15253: gap of unknown length
*   15254 16448: contig of 1195 bp in length
*   16449 16548: gap of unknown length
*   16549 17818: contig of 1270 bp in length
*   17819 17918: gap of unknown length
*   17919 18927: contig of 1009 bp in length
*   18928 19027: gap of unknown length
*   19028 20664: contig of 1637 bp in length
*   20665 20764: gap of unknown length
*   20765 22049: contig of 1285 bp in length
*   22050 22149: gap of unknown length
*   22150 23234: contig of 1085 bp in length
*   23235 23334: gap of unknown length
*   23335 24782: contig of 1448 bp in length
*   24783 24882: gap of unknown length
*   24883 26234: contig of 1352 bp in length
*   26235 26334: gap of unknown length
*   26335 27817: contig of 1483 bp in length
*   27818 27917: gap of unknown length
*   27918 29080: contig of 1163 bp in length
*   29081 29180: gap of unknown length
*   29181 30424: contig of 1244 bp in length
*   30425 30524: gap of unknown length
*   30525 31603: contig of 1079 bp in length
*   31604 31703: gap of unknown length
*   31704 32871: contig of 1168 bp in length

*   32872 32971: gap of unknown length
*   32972 34124: contig of 1153 bp in length
*   34125 34224: gap of unknown length
*   34225 35515: contig of 1291 bp in length
*   35516 35615: gap of unknown length
*   35616 37363: contig of 1748 bp in length
*   37364 37463: gap of unknown length
*   37464 38798: contig of 1335 bp in length
*   38799 38898: gap of unknown length
*   38899 40166: contig of 1268 bp in length
*   40167 40266: gap of unknown length
*   40267 41475: contig of 1209 bp in length
*   41476 41575: gap of unknown length
*   41576 42817: contig of 1242 bp in length
*   42818 42917: gap of unknown length
*   42918 43947: contig of 1030 bp in length
*   43948 44047: gap of unknown length
*   44048 45754: contig of 1707 bp in length
*   45755 45854: gap of unknown length
*   45855 47142: contig of 1288 bp in length
*   47143 47242: gap of unknown length
*   47243 48604: contig of 1362 bp in length
*   48605 48704: gap of unknown length
*   48705 50410: contig of 1706 bp in length
*   50411 50510: gap of unknown length
*   50511 51911: contig of 1401 bp in length
*   51912 52011: gap of unknown length
*   52012 53438: contig of 1427 bp in length
*   53439 53538: gap of unknown length
*   53539 55110: contig of 1572 bp in length
*   55111 55210: gap of unknown length
*   55211 56745: contig of 1535 bp in length
*   56746 56845: gap of unknown length
*   56846 58769: contig of 1924 bp in length
*   58770 58869: gap of unknown length
*   58870 60886: contig of 2017 bp in length
*   60887 60986: gap of unknown length
*   60987 62876: contig of 1890 bp in length
*   62877 62976: gap of unknown length
*   62977 64982: contig of 2006 bp in length
*   64983 65082: gap of unknown length
*   65083 67625: contig of 2543 bp in length
*   67626 67725: gap of unknown length
*   67726 70484: contig of 2759 bp in length
*   70485 70584: gap of unknown length
*   70585 73320: contig of 2736 bp in length
*   73321 73420: gap of unknown length
*   73421 76087: contig of 2667 bp in length
*   76088 76187: gap of unknown length
*   76188 79199: contig of 3012 bp in length
*   79200 79299: gap of unknown length
*   79300 82479: contig of 3180 bp in length.

FEATURES             Location/Qualifiers
     source          1..82479
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-159K5"
BASE COUNT    19500 a  19368 c  18006 g  19214 t     6391 others
ORIGIN

Query Match           11.0%;     Score 20;    DB 2;   Length 82479;
Best Local Similarity  100.0%;   Pred. No. 3.1;
Matches   20;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          93 GCAAGAGGCCCTCCAAGCCA 112
               ||||||||||||||||||||
Db       13464 GCAAGAGGCCCTCCAAGCCA 13483

RESULT 23
AC108558/c
LOCUS       AC108558       169873 bp    DNA    linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-116N1, *** SEQUENCING IN PROGRESS
```

```
ACCESSION   AC108558
VERSION     AC108558.3  GI:21731906
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 169873)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Binage,k., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,R.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,J., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,E., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,E., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169873)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 169873)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 11, 2002 this sequence version replaced gi:1884631l.
            -------------- Genome Center
               Center: Baylor College of Medicine
               Center code: BCM
               Web site: http://www.hgsc.bcm.tmc.edu/
               Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
               Center project name: GLFA
               Center clone name: CH230-116N1
            -------------- Summary Statistics
               Sequencing vector: Plasmid;
               Chemistry: Dye-terminator Big Dye: 100% of reads
               Assembly program: Phrap; version 0.990329
               Consensus quality: 105867 bases at least Q40
               Consensus quality: 111905 bases at least Q30
               Consensus quality: 118103 bases at least Q20
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 66 contigs. The true order of the pieces
            *   is not known and their order in this sequence record is
            *   arbitrary. Gaps between the contigs are represented as
            *   runs of N, but the exact sizes of the gaps are unknown.
            *   This record will be updated with the finished sequence
            *   as soon as it is available and the accession number will
            *   be preserved.
            *
            *          1    1689: contig of 1689 bp in length
            *       1690    1789: gap of unknown length
            *       1790    2950: contig of 1161 bp in length
            *       2951    3050: gap of unknown length
            *       3051    4112: contig of 1062 bp in length
            *       4113    4212: gap of unknown length
            *       4213    5570: contig of 1358 bp in length
            *       5571    5670: gap of unknown length
            *       5671    7141: contig of 1471 bp in length
            *       7142    7241: gap of unknown length
            *       7242    8335: contig of 1094 bp in length
            *       8336    8435: gap of unknown length
            *       8436    9972: contig of 1537 bp in length
            *       9973   10072: gap of unknown length
            *      10073   11551: contig of 1479 bp in length
            *      11552   11651: gap of unknown length
            *      11652   12898: contig of 1247 bp in length
            *      12899   12998: gap of unknown length
            *      12999   14075: contig of 1077 bp in length
            *      14076   14175: gap of unknown length
            *      14176   15203: contig of 1028 bp in length
            *      15204   15303: gap of unknown length
            *      15304   17164: contig of 1861 bp in length
            *      17165   17264: gap of unknown length
            *      17265   18487: contig of 1223 bp in length
            *      18488   18587: gap of unknown length
            *      18588   19671: contig of 1084 bp in length
            *      19672   19771: gap of unknown length
            *      19772   20997: contig of 1226 bp in length
            *      20998   21097: gap of unknown length
            *      21098   22134: contig of 1037 bp in length
            *      22135   22234: gap of unknown length
            *      22235   23386: contig of 1152 bp in length
            *      23387   23486: gap of unknown length
            *      23487   25042: contig of 1556 bp in length
            *      25043   25142: gap of unknown length
            *      25143   27126: contig of 1984 bp in length
            *      27127   27226: gap of unknown length
            *      27227   28324: contig of 1098 bp in length
            *      28325   28424: gap of unknown length
            *      28425   29676: contig of 1252 bp in length
            *      29677   29776: gap of unknown length
            *      29777   31328: contig of 1552 bp in length
            *      31329   31428: gap of unknown length
            *      31429   32526: contig of 1098 bp in length
            *      32527   32626: gap of unknown length
            *      32627   34712: contig of 2086 bp in length
            *      34713   34812: gap of unknown length
            *      34813   36807: contig of 1995 bp in length
            *      36808   36907: gap of unknown length
            *      36908   38969: contig of 2062 bp in length
            *      38970   39069: gap of unknown length
            *      39070   40674: contig of 1605 bp in length
            *      40675   40774: gap of unknown length
            *      40775   42992: contig of 2218 bp in length
```

```
*  42993    43092: gap of unknown length
*  43093    44828: contig of 1736 bp in length
*  44829    44928: gap of unknown length
*  44929    46720: contig of 1792 bp in length
*  46721    46820: gap of unknown length
*  46821    47970: contig of 1150 bp in length
*  47971    48070: gap of unknown length
*  48071    49306: contig of 1236 bp in length
*  49307    49406: gap of unknown length
*  49407    51821: contig of 2415 bp in length
*  51822    51921: gap of unknown length
*  51922    54480: contig of 2559 bp in length
*  54481    54580: gap of unknown length
*  54581    56119: contig of 1539 bp in length
*  56120    56219: gap of unknown length
*  56220    58340: contig of 2121 bp in length
*  58341    58440: gap of unknown length
*  58441    60958: contig of 2518 bp in length
*  60959    61058: gap of unknown length
*  61059    62854: contig of 1796 bp in length
*  62855    62954: gap of unknown length
*  62955    64911: contig of 1957 bp in length
*  64912    65011: gap of unknown length
*  65012    67730: contig of 2719 bp in length
*  67731    67830: gap of unknown length
*  67831    70549: contig of 2719 bp in length
*  70550    70649: gap of unknown length
*  70650    72782: contig of 2133 bp in length
*  72783    72882: gap of unknown length
*  72883    74520: contig of 1638 bp in length
*  74521    74620: gap of unknown length
*  74621    78100: contig of 3480 bp in length
*  78101    78200: gap of unknown length
*  78201    79834: contig of 1634 bp in length
*  79835    79934: gap of unknown length
*  79935    82788: contig of 2854 bp in length
*  82789    82888: gap of unknown length
*  82889    85471: contig of 2583 bp in length
*  85472    85571: gap of unknown length
*  85572    88315: contig of 2744 bp in length
*  88316    88415: gap of unknown length
*  88416    90401: contig of 1986 bp in length
*  90402    90501: gap of unknown length
*  90502    93818: contig of 3317 bp in length
*  93819    93918: gap of unknown length
*  93919    95422: contig of 1504 bp in length
*  95423    95522: gap of unknown length
*  95523    98705: contig of 3183 bp in length
*  98706    98805: gap of unknown length
*  98806   101245: contig of 2440 bp in length
* 101246   101345: gap of unknown length
* 101346   104474: contig of 3129 bp in length
```

```
Query Match             11.0%;  Score 20;  DB 2;  Length 169873;
Best Local Similarity   100.0%;  Pred. No. 3.1;
Matches  20;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    144 CCTGTACCTCATGGTGATGA 163
          ||||||||||||||||||||
Db  90544 CCTGTACCTCATGGTGATGA 90525
```

```
RESULT 24
AB050510
LOCUS       AB050510     1665 bp    mRNA    linear   PRI 27-OCT-2000
DEFINITION  Macaca fascicularis brain cDNA, clone:QnpA-13932.
ACCESSION   AB050510
VERSION     AB050510.1  GI:11041472
KEYWORDS    fis (full insert sequence).
SOURCE      Macaca fascicularis adult male brain parietal lobe cDNA to mRNA,
            clone_lib:macaque brain cDNA library QnpA clone:QnpA-13932.
            Macaca fascicularis
ORGANISM    Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1665)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2000) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab host:  TOP10
            Vector:   pME18S-FL3 (Acc.No. AB009864)
            R. Site1:   DraIII (CACTGTGTG)
            R. Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR. The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert. Libraries
            were constructed by Sugano et al.(University of Tokyo, Institute of
            Medical Science). Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1..1665
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QnpA-13932"
                     /sex="male"
                     /tissue_type="brain parietal lobe"
                     /clone_lib="macaque brain cDNA library QnpA"
                     /dev_stage="adult"
     CDS             487..1044
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB17278.1"
                     /db_xref="GI:11041473"
                     /translation="MRRPVRMRQVVCRLNRSDSDSSTLSKKPPFVRNSLERRSVRMKR
                     PSSVKSLRSERLIRTSLDLEIDLQATRTWHSQLTQBISVLKELKEQLEQAKSHGEKEL
                     PQWLREDERFRLLLRMLEKRQMDQAEHKGELQTDKMMRAAAKDVHRLRGQSCKEPPEV
                     QSFREKMAFFTRPRMNIPALSADDV"
BASE COUNT      512 a   371 c   448 g   332 t     2 others
ORIGIN
```

```
Query Match             10.5%;  Score 19;  DB 9;  Pred. No. 13;  Length 1665;
Best Local Similarity   100.0%;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    15 CACACAGACGCTGGAAGAC 33
         ||||||||||||||||||
Db   458 CACACAGACGCTGGAAGAC 476
```

```
RESULT 25
AB023033/c
LOCUS       AB023033    77129 bp    DNA    linear   PLN 27-DEC-2000
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K6M13.
ACCESSION   AB023033 BA000015
VERSION     AB023033.1  GI:4220632
KEYWORDS    .
SOURCE      Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui TAC
            clone:K6M13.
            Arabidopsis thaliana
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

```
REFERENCE   1  (sites)
AUTHORS     Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kotani,H.
            and Tabata,S.
TITLE       Structural analysis of Arabidopsis thaliana chromosome 5. X.
            Sequence features of the regions of 3,076,755 bp covered by sixty
            P1 and TAC clones
JOURNAL     DNA Res. 7 (1), 31-63 (2000)
MEDLINE     20181125
REFERENCE   2  (bases 1 to 77129)
AUTHORS     Nakamura,Y.
TITLE       Direct Submission
JOURNAL     Submitted (01-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=K6M13
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers. Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is K7J8 and the 3' clone is MNI5.
FEATURES             Location/Qualifiers
     source          1..77129
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="K6M13"
                     /clone_lib="Mitsui TAC"
     CDS             5957..6439
                     /note="gene_id:K6M13.2"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="NaCl-inducible Ca2+-binding protein-like;
                     calmodulin-like"
                     /protein_id="BAB10761.1"
                     /db_xref="GI:10177615"
                     /translation="MRFVPYLYEIERLWLSLVNYLSPRKNKNRRCGEIDKIRMVKKGG
                     TKRKIAIETIQKSDYLRVTCTKRREGLFSKASQLCLLSDAQIAILATPPTSESNISFY
                     SFGHSSVDAVVSSFLSGQRCVPLQEDTKEMREDVAICLSRTNLGLGFWNNESLNKSE
                     NPQEISDAINSMLTLLSNLKELSGEALVNDHKOLKKNERSDVVLQHGTQYETLNPNS
                     NTTICCVPDELPANSNEIVGISPNPLIMLEKKKSQIEEKFEKEWQVSVTRIENEATS
                     SYAKRRRSI"
     CDS             complement(join(12701..13029,13113..13227,13318..13475,
                     13545..13756,13904,13988..14138,14676..14830,
                     14919..15023,15121..15351))
                     /note="gene_id:K6M13.4"
```

```
                     /codon_start=1
                     /evidence=not_experimental
                     /product="SRP54 (signal recognition particle 54 KDa)
                     protein"
                     /protein_id="BAB10763.1"
                     /db_xref="GI:10177616"
                     /translation="MVLAELGRIMSAIQKMNNVTIIDEKALNDCLNEITRALLQSDV
                     SFPLVKEMQTNIKKIVNLEDLAAGHNKRRIIEQAIFSELCKMLDPGKSAFAPKKAKPS
                     VVMFVGLQGEVLEKPQLVPSMLIIIRRKGYKPALVCADTFRAGAFDQLKQNATKSKIP
                     YYGSYTGSDPVKIAVEGVDRFKKENCDLIIVDTSGRHKQQASLFEEMRQISEATKPDL
                     VIFVMDSSIGQTAFEQARAFKQTVAGVAVIITKMDGHAKGGGTLSAVAATKSPVIFIG
                     TGEHMDEFEVFDAKPFVSRLLGNGDMSGFVNKLQEVVPKDQQPELLEMLSHGKFTLRI
                     MYDQFQNMLNMGPLKEVFSMLPGMRAEMMPEGHEKESQAKIKRYMTMMDSMTNEELDS
                     SNPKVFNESRIMRIARGSGRIVREVMELEEYKRLTMWGKVKGLKNLEKGNMSALSK
                     NKNSQQLSKVLPAQMLKQIGGMSGLQSLMKQMGSGKDLMGMPFGGRDE"
                     join(15874..16019,16186..16244,16326..16397,16993..17112,
                     17201..17302,17512..17601)
                     /codon_start=1
     CDS
                     /evidence=not_experimental
                     /product="von Hippel-Lindau binding protein (VHL binding
                     protein; VBP) like"
                     /protein_id="BAB10764.1"
                     /db_xref="GI:10177617"
                     /translation="MSSSPSGSGSDLTERRGIPAAKFIQDVETYLSQSGLDPNSALA
                     FHQERLQQYKVVEMKLLAQQRDLQAKIPDIEKCLEVVATLEAKKGTGEALLADFEVSE
                     GIYSRACIEDTDSVCLWLGANVMLEYSCEEASALLKNNLENAKASLEVLVADLQF_RD
                     QVVTVQTVIARVYNWDVHQRRVKQVVTPTAIAVADS"
                     join(26035..26612,26693..26836,27535..28012)
                     /note="contains similarity to WRKY-type DNA-binding
     CDS
                     protein
                     gene_id:K6M13.6"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB10765.1"
                     /db_xref="GI:10177618"
                     /translation="MEKKKEEDHHQQQQQQKEIKNTETKIEQEQEQKOEISQAS
                     SSNMANLVTSSDHHPLELAGNLSSIPDTSSLPFPYSYFEDHSSNNPNSFLDLLRQDH
                     QFASSNSSSFSDFAPFLPNNNNNTSFFTDLPLPQAESSEVNTTPSPNSTSVSSVSSS
                     NEAANDNNSGKEVTVKDQEEGDQQQEQGTKFQLKAKKKNQKKAREARFAFLTKSDID
                     NLDDGYRWRKYGQKAVKNSPYPRSYYRCTTVGCGVKRVERSSDDPSIVMTYYEGQHT
                     HPFPMTPRGHIGMLTSPILDHGATTASSSSFSIPQPRYLLTOHHQPYNMYNNNSLSMI
                     NRRSSDGTFVNPGPSSSSFPGFGYDMSQASTSTSSSIRDHGLLQDILPSQIRSDTINTQ
                     TNEENKK"
     snRNA           complement(30191..30353)
                     /product="U1 snRNA"
                     /note="gene_id:K6M13.7"
     CDS             join(36079..36483,36574..36642,36719..36820,36894..36962,
                     37053..37175,37284..37334,37442..37548,37651..37722,
                     37871..37922,38062..38690,38820..39029,39120..39240,
                     39928..39387)
                     /note="contains similarity to unknown protein
                     dbj|BAA91655.1
                     gene_id:K6M13.8"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB10766.1"
                     /db_xref="GI:10177619"
                     /translation="MDFDDDDKPKEVTKTRRFAPGRAGKSKPKPEPTADKPVQPPP
                     QSQTESVSKTEHDVDAKFVGTKVETEVCNGSVKMEIDSKVDKEPEIMETELMEDDQL
                     PLQEEKEEEEEDDVVVREIDVFFKPSIDANTQLYVLQYPLRPSWRPYEMDRCEEVR
                     VNPSTSQVEIDLSMDVHSKNYDSNFGLNMTKQTLKTTWKQPPTLDYAVGVLSGDKLHL
                     NPVHAVAQLRPSMQSLSSDKKKQEESTEESVTSKKQNGVQQASTDQKPINEETWV
                     SLKYHGLQSEYCSRYLNGMMAANGNSSIDFNMSPGTYINELCRGGSSRNSESKETLKRV
                     LLSLPKERVQKLLCEGSPLIRYSVLKHYAPEFSDEDFLGALQEYGRLVQGLWTPKTR
                     LLKLDGPVEAARDYVLSLFSQNTTIKYSEVEATGDKMPLMERMLTEFAKERHVLKDW
                     KFKEPTDVSFIKSYPEIVKEQDIFWTDKRENLKSRITAQGGKSRADKRRNVVGTSSSV
                     TVKPEVPTLSDKGGSSKNTIHRVVTQEMPEELKKALPKALKKVFQTHKVRYET:CQ
                     GLRDLAVSTSNNPKADSGMAVNVALAVDAYQGLEDVINGVATNIHGSYVSISSPDHP
                     EYDSLREVVISLLTGSPPGTKLMKAEVPAAGRTKLEREITNNEYIKVMHEICETNSSG
                     WVLQKAR"
                     complement(join(39873..40363,40558,40363..40657,40754))
```

```
                    /note="gene id:K6M13.9
                  unknown protein"
                  /codon_start=1
                  /evidence=not_experimental
                  /protein_id="BAB10767.1"
                  /db_xref="GI:10177620"
                  /translation="MGSSEKRSKDIMSDIPTFSAENLQNNLKVIQNSRTFLSIIAGVL
                  AGIGFNGLIGFVFYFVVMLITSVGLMAKAGFSADLYFDSWNRVLFDGFLGLMNQVY
                  AQLENSSLYSLFVSVGFLNNGVDLRMTWCTYSEGTSAPDEDGETAHKIEFGFCSLFI"
CDS               complement(join(41847. .42146,42583. .42685,42768. .42834,
                  42925. .43045,43119. .43218,43281. .43393,43497. .43612,
                  43753. .43821,43926. .44000,44075. .44116,44231. .44325,
                  44553. .44635,44710. .44760,44862. .44919,45012. .45191,
                  45270. .45431,45512. .45747))
                  /note="gene id:K6M13.10"
                  /codon_start=1
                  /evidence=not_experimental
                  /product="phytoene dehydrogenase-like"
                  /protein_id="BAB10768.1"
                  /db_xref="GI:10177621"
                  /translation="MWRRSFSTTLPKKKWDAVVIGGGHNGLTAAAYLARGGLSVAVLER
                  RHVIGGAAVTEIVPGFKFSRCSYLDQGLLRPCIIRELELGRHGLKLLKRSPSSFTPCL
                  DGRYLLLGPDQDLNHSEISKFSKHDAAYPRYEKQLERFCGFMDPLLDSTPPESLQSA
                  SSFNDKLSNKMYKSAFWARCLRQAVSLGHKDMVAFMDLLLAPASKVLNNWFSDVLKA
                  SLATDAVIGSTASVHTPGSGYVLLHHVMGETDGEKGIWSYVEGGMGSVSMAIANAAKE
                  AGAEIFTNAEVSEILTEDSSIVKGVLLADGTRVESSAILSNATPYRTYVELVPTNVLP
                  ENFVSAIKNSDYSSATTKINLAVDKLPQFQCCNTNHSGPGPEHFGTIHIGAESMDEVH
                  SACHDSENGLPSRRPVIEMTIPSTLDNTISPPGKHVINLFIQYTPYKPSDGSWEDPTY
                  RRAFAQRCFKLIDEYAPGFSSSIISYDMLTPPDLEREIGLTGGNIFHGAMGLDSLFLM
                  RPVKGWSNYRSPLKGLYLICGSGAHPGGGVMGAPGRNAAHVGTNNHLDLLEKMNLRVAS
                  EYDDMGDVAAGLRVFAEQMKSKSGGLDEFVGQMDAIEKQVSBFEAVISVLDRYVSVLE
                  SKIRAEYRHPHHQRRSNDSVVTD"
CDS               complement(46106. .46930)
                  /note="gene id:K6M13.11
                  pir||T08406"
                  similar to unknown protein"
                  /codon_start=1
                  /evidence=not_experimental
                  /protein_id="BAB10769.1"
                  /db_xref="GI:10177622"
                  /translation="MITILPLRDDEENDTYEAKMLIGENGLDGESLPPTPLRDGALE
                  LQKYNISSIESTVVIRELTSQGLSFQLWPAASTFVTLLDNYRRDPSKSPLTATLLSLK
                  KPSPLNLELGSGTGLVGIAAAITLSANVTVTDLPHVLDNLFNAEANAEIVERFGGK
                  VNVAPLRWGEADDVEVLGQNVDLILASDVVYHDHLYEPLLKTLRLMOLEGKRLIFLMA"
```

```
Query Match       10.5%;   Score 19;   DB 8;   Length 77129;
Best Local Similarity   100.0%;   Pred. No. 12;
Matches    19;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy           47  TTTTTATTACTTATATGGA  65
                 |||||||||||||||||||
Db        54007  TTTTTATTACTTATATGGA  53989
```

```
RESULT 26
AC113931/c
LOCUS         AC113931               100466 bp    DNA     linear   PRI 05-MAR-2002
DEFINITION    Homo sapiens chromosome 16 clone CTD-2330P21, complete sequence.
ACCESSION     AC113931
VERSION       AC113931.1  GI:19111663
KEYWORDS      HTG.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 100466)
AUTHORS       DOE Joint Genome Institute.
TITLE         Sequencing of Human Chromosome 16
JOURNAL       Unpublished
REFERENCE     2  (bases 1 to 100466)
AUTHORS       DOE Joint Genome Institute.
TITLE         Direct Submission
JOURNAL       Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint
```

```
COMMENT           Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
                  Sequence Quality Assessment:
                  This entry has been annotated with sequence quality
                  estimates computed by the Phrap assembly program.
                  All manually edited bases have been reduced to quality zero.
                  Quality levels above 40 are expected to have less than
                  1 error in 10,000 bp.
                  Base-by-base quality values are not generally visible from the
                  GenBank flat file format but are available as part
                  of this entry's ASN.1 file.
                  --------------------------
                  Sequence Quality Assessment:
                  This entry has been annotated with sequence quality
                  estimates computed by the Phrap assembly program.
                  All manually edited bases have been reduced to quality zero.
                  Quality levels above 40 are expected to have less than
                  1 error in 10,000 bp.
                  Base-by-base quality values are not generally visible from the
                  GenBank flat file format but are available as part
                  of this entry's ASN.1 file.
                  --------------------------

FEATURES          Location/Qualifiers
     source       1..100466
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="16"
                  /clone="CTD-2330P21"

BASE COUNT   29749 a  20882 c  21101 g  28734 t
ORIGIN

Query Match       10.5%;   Score 19;   DB 9;   Length 100466;
Best Local Similarity   100.0%;   Pred. No. 12;
Matches    19;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy          128  ACTTCTACTATGTCATCCT  146
                 |||||||||||||||||||
Db         2809  ACTTCTACTATGTCATCCT  2791
```

```
RESULT 27
AL359851
LOCUS         AL359851               109920 bp    DNA     linear   PRI 09-MAR-2001
DEFINITION    Human DNA sequence from clone RP13-192B19 on chromosome Xq24-25,
              complete sequence.
ACCESSION     AL359851
VERSION       AL359851.19  GI:13274766
KEYWORDS      HTG.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 109920)
AUTHORS       Blakey,S.
TITLE         Direct Submission
JOURNAL       Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
              requests: clonerequest@sanger.ac.uk
COMMENT       On Mar 12, 2001 this sequence version replaced gi:12956948.
              During sequence assembly data is compared from overlapping clones.
              Where differences are found these are annotated as variations
              together with a note of the overlapping clone name. Note that the
              variation annotation may not be found in the sequence submission
              corresponding to the overlapping clone, as we submit sequences with
              only a small overlap as described above.
              This sequence was finished as follows unless otherwise noted: all
              regions were either double-stranded or sequenced with an alternate
              chemistry or covered by high quality data (i.e. phred quality >=
              30); an attempt was made to resolve all sequencing problems, such
              as compressions and repeats; all regions were covered by at least
              one plasmid subclone or more than one M13 subclone; and the
              assembly was confirmed by restriction digest. The following
              abbreviations are used to associate primary accession numbers given
              in the feature table with their source databases: Em:, EMBL; Sw:;
```

```
        SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
        database can be found at
        http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
        was generated from part of bacterial clone contigs of human
        chromosome X, constructed by the Sanger Centre Chromosome X Mapping
        Group. Further information can be found at
        http://www.sanger.ac.uk/HGP/ChrX
        RP13-192B19 is from the library RPCI-13.1 constructed by the group
        of Pieter de Jong. For further details see
        http://www.chori.org/bacpac/home.htm
        VECTOR: pBACe3.6
        IMPORTANT: This sequence is not the entire insert of clone
        RP13-192B19 It may be shorter because we sequence overlapping
        sections only once, except for a 100 base overlap.
        The true left end of clone RP13-192B19 is at 1 in this sequence.
        The true left end of clone RP11-417D4 is at 109821 in this
        sequence.

FEATURES             Location/Qualifiers
     source          1..109920
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="q24-25"
                     /clone="RP13-192B19"
                     /clone_lib="RPCI-13.1"
     repeat_region   275..468
                     /note="97 copies 2 mer tt 57% conserved"
     repeat_region   1425..3361
                     /note="L1PBa repeat: matches -942..1930 of consensus"
     repeat_region   3262..3458
                     /note="L1PBa repeat: matches -1540..-1343 of consensus"
     repeat_region   3678..4769
                     /note="MLT1A1-internal repeat: matches 1..1063 of
                     consensus"
     repeat_region   4770..5156
                     /note="MLT1A1 repeat: matches 9..365 of consensus"
     repeat_region   5158..5958
                     /note="AluSx repeat: matches 1..305 of consensus"
     repeat_region   6270..6579
                     /note="L1M4 repeat: matches 3554..4429 of consensus"
     repeat_region   6603..6774
                     /note="AluJo repeat: matches 1..304 of consensus"
     repeat_region   7946..9909
                     /note="L1ME3 repeat: matches 5970..6141 of consensus"
     repeat_region   9072..9376
                     /note="L1ME3 repeat: matches 4890..5984 of consensus"
     repeat_region   9919..10004
                     /note="AluSx repeat: matches 1..305 of consensus"
     repeat_region   10009..10042
                     /note="L1MC4 repeat: matches 7863..7951 of consensus"
     repeat_region   10105..10493
                     /note="17 copies 2 mer gt 88% conserved"
     repeat_region   10494..11031
                     /note="MLT1C repeat: matches 6..432 of consensus"
     repeat_region   11032..11069
                     /note="MER44B repeat: matches 2..536 of consensus"
     repeat_region   11070..11574
                     /note="MLT1C repeat: matches 432..466 of consensus"
     repeat_region   11669..11758
                     /note="L1MC4 repeat: matches 7251..7787 of consensus"
     repeat_region   11786..11879
                     /note="L1MC5 repeat: matches 7132..7221 of consensus"
     repeat_region   12500..12578
                     /note="47 copies 2 mer ct 68% conserved"
     repeat_region   12594..12694
                     /note="MADE1 repeat: matches 1004..1106 of consensus"
     repeat_region   12842..13027
                     /note="MER52A repeat: matches 5467..5651 of consensus"
     repeat_region   13463..15156
                     /note="L1MC/D repeat: matches 1..1753 of consensus"
     repeat_region   15173..15722
                     /note="MER52A repeat: matches 1..1753 of consensus"
     repeat_region   16395..16539
                     /note="L1M4 repeat: matches 4475..5025 of consensus"

     repeat_region   17571..17975
                     /note="L1MD3 repeat: matches 7588..7734 of consensus"
     repeat_region   19384..19439
                     /note="MER92A repeat: matches 1..412 of consensus"
     repeat_region   19858..19906
                     /note="28 copies 2 mer aa 71% conserved"
     repeat_region   19915..20035
                     /note="L1PA16 repeat: matches 6107..6157 of consensus"
     repeat_region   20036..20342
                     /note="MER2 repeat: matches 106..210 of consensus"
     repeat_region   20988..21430
                     /note="LTR24 repeat: matches 1..443 of consensus"
     repeat_region   21431..27142
                     /note="L1PA3 repeat: matches 8..6146 of consensus"
     repeat_region   27351..27692
                     /note="MSTD repeat: matches 1..343 of consensus"
     repeat_region   27696..27719
                     /note="12 copies 2 mer aa 95% conserved"
     repeat_region   27970..28213
                     /note="HAL1 repeat: matches 258..515 of consensus"
     repeat_region   28263..28568
                     /note="AluSx repeat: matches 1..305 of consensus"
     repeat_region   28584..28731
                     /note="L1MD1 repeat: matches 6073..6218 of consensus"
     repeat_region   28732..28768
                     /note="MER3 repeat: matches 20..56 of consensus"
     repeat_region   28769..29267
                     /note="L1MD1 repeat: matches 5530..6073 of consensus"
     repeat_region   33604..33953
                     /note="THE1A repeat: matches 1..354 of consensus"
     repeat_region   35764..36123
                     /note="MLT1A1 repeat: matches 1..365 of consensus"
     repeat_region   37297..37346
                     /note="25 copies 2 mer at 72% conserved"
     repeat_region   37783..38345
                     /note="L1 repeat: matches 3685..4337 of consensus"
     repeat_region   38347..38508
                     /note="TIGGER1 repeat: matches 2202..2417 of consensus"
     repeat_region   38578..38736
                     /note="TIGGER1 repeat: matches 1..157 of consensus"
     repeat_region   38813..39255
                     /note="MLT2A repeat: matches 1..453 of consensus"
     repeat_region   39294..39327
                     /note="17 copies 2 mer ac 100% conserved"
     repeat_region   39330..39394
                     /note="MER8 repeat: matches 177..236 of consensus"
     repeat_region   39546..39783
                     /note="Tigger2a repeat: matches 15..434 of consensus"
     repeat_region   39854..39942
                     /note="L1ME3 repeat: matches 5616..5704 of consensus"
     repeat_region   39990..40330
                     /note="L1ME3 repeat: matches 5782..6140 of consensus"
     repeat_region   40347..40392
                     /note="23 copies 2 mer ta 78% conserved"
     repeat_region   40499..40878
                     /note="MLT1A1 repeat: matches 1..365 of consensus"
     repeat_region   41130..41335
                     /note="LTR16A repeat: matches 225..438 of consensus"
     repeat_region   41475..41791
                     /note="AluY repeat: matches 1..309 of consensus"
     repeat_region   41793..41915
                     /note="L1ME repeat: matches 5495..5618 of consensus"
     repeat_region   41941..42181
                     /note="L1ME1 repeat: matches 5679..5947 of consensus"
     repeat_region   42679..42843
                     /note="L1PA12 repeat: matches 5857..6018 of consensus"
     repeat_region   42844..42966
                     /note="L1PA13 repeat: matches 6034..6156 of consensus"
     repeat_region   42974..43076
                     /note="MER5A repeat: matches 11..109 of consensus"
     repeat_region   43321..44603
                     /note="MER63C repeat: matches 52..786 of consensus"
     repeat_region   44766..44960
                     /note="L1MC3 repeat: matches 7525..7735 of consensus"
```

```
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the library RPCI-4 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP4-569D19 It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone RP4-569D19 is at 1 in this sequence. The
true right end of clone RP5-824I19 is at 31286 in this sequence.
The start of this sequence overlaps with sequence AL009049.

FEATURES           Location/Qualifiers
     source        1..114771
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /chromosome="22"
                   /map="q13.1"
                   /clone="RP4-569D19"
                   /clone_lib="RPCI-4"
     repeat_region 805..1011
                   /note="MIR repeat: matches 47. .258 of consensus"
     repeat_region 1012..1253
                   /note="AluSc repeat: matches 61. .304 of consensus"
     repeat_region 1466..1613
                   /note="MIR repeat: matches 36. .237 of consensus"
     repeat_region 1643..2088
                   /note="L2 repeat: matches 2487. .2621 of consensus"
     repeat_region 2302..2611
                   /note="L2 repeat: matches 1929. .2417 of consensus"
     repeat_region 2748..3041
                   /note="AluSg repeat: matches 1. .309 of consensus"
     repeat_region 3546..3871
                   /note="L2 repeat: matches 2470. .2750 of consensus"
     repeat_region 3918..4117
                   /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region 4206..4486
                   /note="MIR repeat: matches 36. .237 of consensus"
     repeat_region 4747..4858
                   /note="AluJb repeat: matches 1. .282 of consensus"
     repeat_region 4882..4959
                   /note="MIR repeat: matches 47. .163 of consensus"
     repeat_region 4966..5077
                   /note="L2 repeat: matches 2625. .2704 of consensus"
     repeat_region 5693..5990
                   /note="MIR repeat: matches 145. .256 of consensus"
     repeat_region 5999..6140
                   /note="AluY repeat: matches 1. .300 of consensus"
     repeat_region 6147..6202
                   /note="AluSq/x repeat: matches 1. .136 of consensus"
     repeat_region 6453..7182
                   /note="MIR repeat: matches 151. .205 of consensus"
     repeat_region 7294..7605
                   /note="L1MA4 repeat: matches 4129. .4859 of consensus"
     repeat_region 7786..8087
                   /note="AluSx repeat: matches 1. .312 of consensus"
     misc_feature  8277..8661
                   /note="AluSg repeat: matches 6. .307 of consensus"
     misc_feature  8748..8909
                   /note="match: STS: Em:R85597"
     repeat_region 8910..8965
                   /note="AluYa5 repeat: matches 150. .311 of consensus"
     misc_feature  complement(9476. .9902)
                   /note="28 copies 2 mer at 100 conserved"
     misc_feature  complement(9546. .9904)
                   /note="match: GSS: Em:AQ222044"
     misc_feature  9891. .10367
                   /note="match: GSS: Em:B57571"
     repeat_region 10091. .10506
                   /note="match: GSS: Em:AQ061599"
     repeat_region 10508. .10592
                   /note="L1MA4 repeat: matches 4845. .5273 of consensus"
                   /note="AluJ/FLAM repeat: matches 1. .88 of consensus"

     repeat_region 44947..45762
                   /note="L1M4 repeat: matches 4543. .5384 of consensus"
     repeat_region 45753..45860
                   /note="L1M4 repeat: matches 2467. .2575 of consensus"
     repeat_region 45850..46590
                   /note="L1MCb repeat: matches 104. .836 of consensus"
     repeat_region 46717..47044
                   /note="L1MB4 repeat: matches 5835. .6163 of consensus"
     repeat_region 47047..47440
                   /note="MER61B repeat: matches 1. .396 of consensus"
     repeat_region 47452..48284
                   /note="MER61-internal repeat: matches 1. .2371 of
                   consensus"
     repeat_region 48284..48481
                   /note="HUERS-P3b repeat: matches 3164. .3360 of consensus"
     repeat_region 48364..51335
                   /note="LTR25-internal repeat: matches 3736. .6713 of
                   consensus"
     repeat_region 48663..52499
                   /note="MER61-internal repeat: matches 742. .4550 of
                   consensus"
     repeat_region 52506..52906

Query Match             10.5%;  Score 19;  DB 9;  Length 109920;
Best Local Similarity  100.0%;  Pred. No. 12;
Matches   19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      111 CAAAGTTGATGCTGAGAAC 129
            |||||||||||||||||||
Db    80309 CAAAGTTGATGCTGAGAAC 80327

RESULT 28
HS569D19/c
LOCUS       HS569D19      114771 bp    DNA        linear    PRI 12-DEC-1999
DEFINITION  Human DNA sequence from clone RP4-569D19 on chromosome 22q13.1
            Contains 3' part of a novel gene similar to mouse Ras,
            Dexamethasone-induced 1 (Ras-related protein, RASD1, DEXRAS1) and
            the last exon of the MB gene for Myoglobin, ESTs, STSs and GSSs,
            complete sequence.
ACCESSION   AL022334
VERSION     AL022334.1  GI:3947839
KEYWORDS    HTG; CpG Island; Dexamethasone-induced; DEXRAS1; Myoglobin; RASD1.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114771)
AUTHORS     Williams,S.
TITLE       Direct Submission
JOURNAL     Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequests@sanger.ac.uk
COMMENT     On Dec 2, 1998 this sequence version replaced gi:3445463.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
```

repeat_region   10593..11608
                /note="L1MA4 repeat: matches 5265..6298 of consensus"
repeat_region   11938..12082
                /note="L1ME1 repeat: matches 6010..6156 of consensus"
variation   12275..12277
                /note="clone RP5-824I19; atg in this entry; substitution"
                /replace="acg"
repeat_region   12344..12784
                /note="L1ME1 repeat: matches 5576..6016 of consensus"
repeat_region   13431..13736
                /note="AluSx repeat: matches 1..306 of consensus"
repeat_region   13792..14127
                /note="LTR9 repeat: matches 10..290 of consensus"
repeat_region   14128..14415
                /note="AluSx repeat: matches 1..289 of consensus"
repeat_region   14416..14720
                /note="LTR9 repeat: matches 290..612 of consensus"
repeat_region   14909..15033
                /note="FLAM_C repeat: matches 2..126 of consensus"
repeat_region   15083..15640
                /note="L2 repeat: matches 2157..2743 of consensus"
repeat_region   15617..15787
                /note="MIR repeat: matches 85..261 of consensus"
repeat_region   16487..16510
                /note="12 copies 2 mer ct 100 conserved"
misc_feature   16658..16913
                /note="match: GSS: Em:B13882"
repeat_region   17863..18145
                /note="LTR16A repeat: matches 172..449 of consensus"
repeat_region   18408..18441
                /note="17 copies 2 mer gt 97 conserved"
variation   18417..18422
                /note="clone RP5-824I19; tgtgtg in this entry; insertion"
                /replace="tg"
repeat_region   18686..19065
                /note="MLT1J repeat: matches 1..398 of consensus"
repeat_region   19286..19459
                /note="MER5A repeat: matches 16..188 of consensus"
repeat_region   19562..19687
                /note="L2 repeat: matches 2619..2750 of consensus"
repeat_region   19703..19851
                /note="AluJo/FRAM repeat: matches 164..285 of consensus"
variation   19877..19914
                /note="19 copies 2 mer gt 100 conserved"
variation   19877..19880
                /note="clone RP5-824I19; gtgt in this entry; insertion"
                /replace="gt"
repeat_region   19917..20520
                /note="L2 repeat: matches 1877..2592 of consensus"
repeat_region   20599..20638
                /note="68 copies 2 mer tc 61 conserved"
repeat_region   20584..22627
                /note="L2 repeat: matches 2584..2624 of consensus"
repeat_region   21196..21503
                /note="22 copies 2 mer ct 81 conserved"
repeat_region   22629..22940
                /note="AluJo repeat: matches 1..305 of consensus"
repeat_region   21798..21922
                /note="FLAM_C repeat: matches 1..125 of consensus"
repeat_region   22058..22365
                /note="AluSq repeat: matches 1..312 of consensus"
repeat_region   22448..22583
                /note="AluSx repeat: matches 1..312 of consensus"
repeat_region   23459..23757
                /note="AluJb repeat: matches 1..312 of consensus"
repeat_region   23793..24045
                /note="AluSq repeat: matches 1..292 of consensus"
repeat_region   24189..24357
                /note="MIR repeat: matches 7..262 of consensus"
repeat_region   26076..26159
                /note="MIR repeat: matches 47..220 of consensus"
repeat_region   26451..26453
                /note="MIR repeat: matches 59..144 of consensus"
variation   26451..26453
                /note="clone RP5-824I19; gcg in this entry; substitution"

repeat_region   26551..26574
                /note="12 copies 2 mer tt 95 conserved"
repeat_region   26837..26996
                /note="MIR repeat: matches 41..193 of consensus"
variation   26955..26957
                /note="clone RP5-824I19; tct in this entry; substitution"
                /replace="ttt"
variation   27020..27022
                /note="clone RP5-824I19; cat in this entry; substitution"
                /replace="cct"
variation   27055..27057
                /note="clone RP5-824I19; atg in this entry; substitution"
                /replace="agg"
variation   27084..27086
                /replace="gtg"

Query Match   10.5%; Score 19; DB 9; Length 11477l;
Best Local Similarity 100.0%; Pred. No. 12;
Matches   19; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   155 TGGTGATGATTGGAATGTT 173
            |||||||||||||||||||
Db   62214 TGGTGATGATTGGAATGTT 62196

RESULT 29
AC097300
LOCUS   AC097300   120531 bp   DNA   linear   HTG 12-JUL-2002
DEFINITION Rattus norvegicus clone CH230-164J5, *** SEQUENCING IN PROGRESS
            ***, 54 unordered pieces.
ACCESSION   AC097300
VERSION   AC097300.3 GI:21728571
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 120531)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Binage,K., Brieva,M., Brown,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,M., Brown,N.P., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,J., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,H., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., McLeod,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,J., Newton,D., Newton,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,J., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,

Wed Jun 11 10:49:33 2003   us-09-550-163-1_copy_80_260.oligo.rge

```
          Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
          Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
          Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
          Weinstock,G. and Gibbs,R.
TITLE     Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 120531)
AUTHORS   Worley,K.C.
TITLE     Direct Submission
JOURNAL   Submitted (14-OCT-2001) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 120531)
AUTHORS   Worley,K.C.
TITLE     Direct Submission
JOURNAL   Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Jul 11, 2002 this sequence version replaced gi:17064442.
          -------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
            Center project name: GIBH
            Center clone name: CH230-164J5
          -------------- Summary Statistics
            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 76061 bases at least Q40
            Consensus quality: 82960 bases at least Q30
            Consensus quality: 89454 bases at least Q20
          * NOTE: Estimated insert size may differ from sequence length
          *     (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 54 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          *        1      1299: contig of 1299 bp in length
          *     1300      1399: gap of unknown length
          *     1400      2459: contig of 1060 bp in length
          *     2460      2559: gap of unknown length
          *     2560      4235: contig of 1676 bp in length
          *     4236      4335: gap of unknown length
          *     4336      5784: contig of 1449 bp in length
          *     5785      5884: gap of unknown length
          *     5885      7133: contig of 1249 bp in length
          *     7134      7233: gap of unknown length
          *     7234      8419: contig of 1186 bp in length
          *     8420      8519: gap of unknown length
          *     8520      9551: contig of 1032 bp in length
          *     9552      9651: gap of unknown length
          *     9652     11069: contig of 1418 bp in length
          *    11070     11169: gap of unknown length
          *    11170     12336: contig of 1167 bp in length
          *    12337     12436: gap of unknown length
          *    12437     13634: contig of 1198 bp in length
          *    13635     13734: gap of unknown length
          *    13735     16230: contig of 2496 bp in length
          *    16231     16330: gap of unknown length
          *    16331     17482: contig of 1152 bp in length
          *    17483     17582: gap of unknown length
          *    17583     19093: contig of 1511 bp in length
          *    19094     19193: gap of unknown length
          *    19194     20672: contig of 1479 bp in length
          *    20673     20772: gap of unknown length
          *    20773     21899: contig of 1127 bp in length
          *    21900     21999: gap of unknown length
          *    22000     23643: contig of 1644 bp in length
          *    23644     23743: gap of unknown length
          *    23744     25283: contig of 1540 bp in length
          *    25284     25383: gap of unknown length
          *    25384     26672: contig of 1289 bp in length
          *    26673     26772: gap of unknown length
          *    26773     27987: contig of 1215 bp in length
          *    27988     28087: gap of unknown length
          *    28088     29237: contig of 1150 bp in length
          *    29238     29337: gap of unknown length
          *    29338     31430: contig of 2093 bp in length
          *    31431     31530: gap of unknown length
          *    31531     32897: contig of 1367 bp in length
          *    32898     32997: gap of unknown length
          *    32998     34633: contig of 1636 bp in length
          *    34634     34733: gap of unknown length
          *    34734     37098: contig of 2365 bp in length
          *    37099     37198: gap of unknown length
          *    37199     38606: contig of 1408 bp in length
          *    38607     38706: gap of unknown length
          *    38707     40214: contig of 1508 bp in length
          *    40215     40314: gap of unknown length
          *    40315     42010: contig of 1696 bp in length
          *    42011     42110: gap of unknown length
          *    42111     44058: contig of 1948 bp in length
          *    44059     44158: gap of unknown length
          *    44159     45667: contig of 1509 bp in length
          *    45668     45767: gap of unknown length
          *    45768     48315: contig of 2548 bp in length
          *    48316     48415: gap of unknown length
          *    48416     50945: contig of 2530 bp in length
          *    50946     51045: gap of unknown length
          *    51046     53236: contig of 2191 bp in length
          *    53237     53336: gap of unknown length
          *    53337     55316: contig of 1980 bp in length
          *    55317     55416: gap of unknown length
          *    55417     57580: contig of 2164 bp in length
          *    57581     57680: gap of unknown length
          *    57681     60289: contig of 2609 bp in length
          *    60290     60389: gap of unknown length
          *    60390     62857: contig of 2468 bp in length
          *    62858     62957: gap of unknown length
          *    62958     65482: contig of 2525 bp in length
          *    65483     65582: gap of unknown length
          *    65583     68725: contig of 3143 bp in length
          *    68726     68825: gap of unknown length
          *    68826     71212: contig of 2387 bp in length
          *    71213     71312: gap of unknown length
          *    71313     73574: contig of 2262 bp in length
          *    73575     73674: gap of unknown length
          *    73675     76601: contig of 2927 bp in length
          *    76602     76701: gap of unknown length
          *    76702     79779: contig of 3078 bp in length
          *    79780     79879: gap of unknown length
          *    79880     83588: contig of 3709 bp in length
          *    83589     83688: gap of unknown length
          *    83689     86943: contig of 3255 bp in length
          *    86944     87043: gap of unknown length
          *    87044     89794: contig of 2751 bp in length
          *    89795     89894: gap of unknown length
          *    89895     92781: contig of 2887 bp in length
          *    92782     92881: gap of unknown length
          *    92882     95811: contig of 2930 bp in length
          *    95812     95911: gap of unknown length
          *    95912     98619: contig of 2708 bp in length
          *    98620     98719: gap of unknown length
          *    98720    101655: contig of 2936 bp in length
          *   101656    101755: gap of unknown length
          *   101756    104324: contig of 2569 bp in length
```

```
*    104325   104424: gap of unknown length
*    104425   107133: contig of 2709 bp in length
*    107134   107233: gap of unknown length
*    107234   110176: contig of 2943 bp in length
*    110177   110276: gap of unknown length
*    110277   115200: contig of 4924 bp in length
*    115201   115300: gap of unknown length
*    115301   120531: contig of 5231 bp in length.

Query Match          10.5%;   Score 19;   DB 2;   Length 120531;
Best Local Similarity 100.0%;  Pred. No. 12;
Matches  19; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy         45 GATTTTTATTACTTATATG 63
              ||||||||||||||||||
Db     109187 GATTTTTATTACTTATATG 109205


RESULT 30
AC016597/c
LOCUS       AC016597           129338 bp    DNA     linear   PRI 20-JUN-2000
DEFINITION  Homo sapiens chromosome 16 clone CTD-204904, complete sequence.
ACCESSION   AC016597
VERSION     AC016597.4  GI:8573019
KEYWORDS    HTG.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 129338)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 129338)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 129338)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jun 20, 2000 this sequence version replaced gi:7690222.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 0.7.
FEATURES             Location/Qualifiers
     source          1..129338
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-204904"
BASE COUNT    34275 a  30831 c  30951 g  33181 t
ORIGIN

Query Match          10.5%;   Score 19;   DB 9;   Length 129338;
Best Local Similarity 100.0%;  Pred. No. 12;
Matches  19; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy        161 TGATTGGAATGTTCTCTTT 179
              |||||||||||||||||||
Db      55078 TGATTGGAATGTTCTCTTT 55060


RESULT 31
AC078804
LOCUS       AC078804           164456 bp    DNA     linear   HTG 19-JUN-2002
DEFINITION  Homo sapiens chromosome 3 clone RP11-489M17, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC078804
VERSION     AC078804.7  GI:21431061
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164456)
AUTHORS     Muzny,D.M., Adams,J.M., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,N.P., Bryant,N.P.,
            Buhay,C., Burch,P., Burket,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,J., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,J., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shoosntari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,C., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 164456)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (04-AUG-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 164456)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (19-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 17, 2002 this sequence version replaced gi:20335729.
            --------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Drafting Center Code: BCM
            Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information
            Center project name: HBTN
            Center clone name: RP11-489M17
            --------------- Summary Statistics
```

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,H., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newtson,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 164995)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (08-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 164995)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Jul 11, 2002 this sequence version replaced gi:17973833.
COMMENT   Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GINP
                Center clone name: CH230-178A23
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 108374 bases at least Q40
                Consensus quality: 114597 bases at least Q30
                Consensus quality: 120549 bases at least Q20
            ----
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 61 contigs. The true order of the pieces
            *   is not known and their order in this sequence record is
            *   arbitrary. Gaps between the contigs are represented as
            *   runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            *   as soon as it is available and the accession number will
            *   be preserved.

Sequencing vector: M13;
Chemistry: Dye-primer Bodipy: 75% of reads
Chemistry: Dye-terminator Big Dye: 25% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 158577 bases at least Q40
Consensus quality: 160839 bases at least Q30
Consensus quality: 161950 bases at least Q20
Estimated insert size: 165132; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length.
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 7 contigs. The true order of the pieces
*   is not known and their order in this sequence record is
*   arbitrary. Gaps between the contigs are represented as
*   runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
*   as soon as it is available and the accession number will
*   be preserved.

FEATURES   Location/Qualifiers
source   1..164456
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="3"
                /clone="RP11-489M17"
BASE COUNT   52500 a 30465 c 29464 g 50085 t   1942 others
ORIGIN

Query Match   10.5%;   Score 19; DB 2; Length 164456;
Best Local Similarity 100.0%;   Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACTTTATCCAATTTCACAC 19
           |||||||||||||||||||
Db 117867 ACTTTATCCAATTTCACAC 117885

                  1 7694: contig of 7694 bp in length
               7695 7794: gap of unknown length
               7795 16159: contig of 8365 bp in length
              16160 16259: gap of unknown length
              16260 31239: contig of 14980 bp in length
              31240 31339: gap of unknown length
              31340 52212: contig of 20873 bp in length
              52213 52312: gap of unknown length
              52313 78975: contig of 26663 bp in length
              78976 79075: gap of unknown length
              79076 118007: contig of 38932 bp in length
             118008 118107: gap of unknown length
             118108 164456: contig of 46349 bp in length

RESULT 32
AC098993/c
LOCUS   AC098993   164995 bp   DNA   linear   HTG 12-JUL-2002
DEFINITION Rattus norvegicus clone CH230-178A23, *** SEQUENCING IN PROGRESS
            ***, 61 unordered pieces.
ACCESSION   AC098993
VERSION   AC098993.4 GI:21729909
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 164995)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

```
*      1   1112: contig of 1112 bp in length
*   1113   1212: gap of unknown length
*   1213   2467: contig of 1255 bp in length
*   2468   2567: gap of unknown length
*   2568   3740: contig of 1173 bp in length
*   3741   3840: gap of unknown length
*   3841   5483: contig of 1643 bp in length
*   5484   5583: gap of unknown length
*   5584   7080: contig of 1497 bp in length
*   7081   7180: gap of unknown length
*   7181   8459: contig of 1279 bp in length
*   8460   8559: gap of unknown length
*   8560  10014: contig of 1455 bp in length
*  10015  10114: gap of unknown length
*  10115  11848: contig of 1734 bp in length
*  11849  11948: gap of unknown length
*  11949  13211: contig of 1263 bp in length
*  13212  13311: gap of unknown length
*  13312  15073: contig of 1762 bp in length
*  15074  15173: gap of unknown length
*  15174  16324: contig of 1151 bp in length
*  16325  16424: gap of unknown length
*  16425  17580: contig of 1156 bp in length
*  17581  17680: gap of unknown length
*  17681  19046: contig of 1366 bp in length
*  19047  19146: gap of unknown length
*  19147  20652: contig of 1506 bp in length
*  20653  20752: gap of unknown length
*  20753  22043: contig of 1291 bp in length
*  22044  22143: gap of unknown length
*  22144  23307: contig of 1164 bp in length
*  23308  23407: gap of unknown length
*  23408  25328: contig of 1921 bp in length
*  25329  25428: gap of unknown length
*  25429  27015: contig of 1587 bp in length
*  27016  27115: gap of unknown length
*  27116  28581: contig of 1466 bp in length
*  28582  28681: gap of unknown length
*  28682  30018: contig of 1337 bp in length
*  30019  30118: gap of unknown length
*  30119  32268: contig of 2150 bp in length
*  32269  32368: gap of unknown length
*  32369  33508: contig of 1140 bp in length
*  33509  33608: gap of unknown length
*  33609  35572: contig of 1964 bp in length
*  35573  35672: gap of unknown length
*  35673  37446: contig of 1774 bp in length
*  37447  37546: gap of unknown length
*  37547  40443: contig of 2897 bp in length
*  40444  40543: gap of unknown length
*  40544  42752: contig of 2209 bp in length
*  42753  42852: gap of unknown length
*  42853  45407: contig of 2555 bp in length
*  45408  45507: gap of unknown length
*  45508  47744: contig of 2237 bp in length
*  47745  47844: gap of unknown length
*  47845  49868: contig of 2024 bp in length
*  49869  49968: gap of unknown length
*  49969  51911: contig of 1943 bp in length
*  51912  52011: gap of unknown length
*  52012  54549: contig of 2538 bp in length
*  54550  54649: gap of unknown length
*  54650  56260: contig of 1611 bp in length
*  56261  56360: gap of unknown length
*  56361  58767: contig of 2407 bp in length
*  58768  58867: gap of unknown length
*  58868  61817: contig of 2950 bp in length
*  61818  61917: gap of unknown length
*  61918  65545: contig of 3628 bp in length
*  65546  65645: gap of unknown length
*  65646  68984: contig of 3339 bp in length
*  68985  69084: gap of unknown length
*  69085  70717: contig of 1633 bp in length
*  70718  70817: gap of unknown length
*  70818  74216: contig of 3399 bp in length
*  74217  74316: gap of unknown length
*  74317  78217: contig of 3901 bp in length
*  78218  78317: gap of unknown length
*  78318  80203: contig of 1886 bp in length
*  80204  80303: gap of unknown length
*  80304  83260: contig of 2957 bp in length
*  83261  83360: gap of unknown length
*  83361  87220: contig of 3860 bp in length
*  87221  87320: gap of unknown length
*  87321  90185: contig of 2865 bp in length
*  90186  90285: gap of unknown length
*  90286  93709: contig of 3424 bp in length
*  93710  93809: gap of unknown length
*  93810  95921: contig of 2112 bp in length
*  95922  96021: gap of unknown length
*  96022  99369: contig of 3348 bp in length
*  99370  99469: gap of unknown length
*  99470 102559: contig of 3090 bp in length
* 102560 102659: gap of unknown length
* 102660 105301: contig of 2642 bp in length
* 105302 105401: gap of unknown length
* 105402 108991: contig of 3590 bp in length
* 108992 109091: gap of unknown length
* 109092 111276: contig of 2185 bp in length
* 111277 111376: gap of unknown length
* 111377 114356: contig of 2980 bp in length
* 114357 114456: gap of unknown length
* 114457 118823: contig of 4367 bp in length
* 118824 118923: gap of unknown length
* 118924 123196: contig of 4273 bp in length
* 123197 123296: gap of unknown length
* 123297 128564: contig of 5268 bp in length


Query Match          10.5%;   Score 19;   DB 2;   Length 164995;
Best Local Similarity  100.0%;   Pred. No. 12;
Matches 19;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy      163 ATTGGAATGTTCTCTTCA 181
            ||||||||||||||||||
Db   161789 ATTGGAATGTTCTCTTCA 161771

RESULT 33
AC092130
LOCUS       AC092130              170520 bp    DNA     linear   PRI 02-OCT-2001
DEFINITION  Homo sapiens chromosome 16 clone RP11-148M9, complete sequence.
ACCESSION   AC092130
VERSION     AC092130.2  GI:15825569
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 170520)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 170520)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 170520)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Oct 2, 2001 this sequence version replaced gi:14522976.
            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
```

```
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.

            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.

FEATURES             Location/Qualifiers
     source          1..170520
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-148M9"

BASE COUNT   50790 a   35640 c   34394 g   49696 t
ORIGIN

Query Match     10.5%;    Score 19;  DB 9;  Length 170520;
Best Local Similarity   100.0%;   Pred. No. 12;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    128  ACTTCTACTATGTCATCCT  146
           |||||||||||||||||||
Db  15604  ACTTCTACTATGTCATCCT  15622

RESULT 34
ACCESSION   AC128107
LOCUS       AC128107           176036 bp    DNA     linear   HTG 19-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-71H18, *** SEQUENCING IN PROGRESS
            ***, 70 unordered pieces.
ACCESSION   AC128107
VERSION     AC128107.1  GI:21908705
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus.
ORGANISM    Rattus norvegicus.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 176036)
AUTHORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,K.L., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,N., Neal,D., Newton,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,H., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,D., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 176036)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center

            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information

            Center project name: GYAG
            Center clone name: CH230-71H18
            -------------- Summary Statistics

            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 109378 bases at least Q40
            Consensus quality: 116692 bases at least Q30
            Consensus quality: 122236 bases at least Q20

          * NOTE: Estimated insert size may differ from sequence length
          *       (see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          *       consists of 70 contigs. The true order of the pieces
          *       is not known and their order in this sequence record is
          *       arbitrary. Gaps between the contigs are represented as
          *       runs of N, but the exact sizes of the gaps are unknown.
          *       This record will be updated with the finished sequence
          *       as soon as it is available and the accession number will
          *       be preserved.

          *       1     1119    contig of 1119 bp in length
          *    1120     1219    gap of unknown length
          *    1220     2255    contig of 1036 bp in length
          *    2256     2355    gap of unknown length
          *    2356     3378    contig of 1023 bp in length
          *    3379     3478    gap of unknown length
          *    3479     5064    contig of 1586 bp in length
          *    5065     5164    gap of unknown length
          *    5165     6492    contig of 1328 bp in length
          *    6493     6592    gap of unknown length
          *    6593     7741    contig of 1149 bp in length
          *    7742     7841    gap of unknown length
          *    7842     9078    contig of 1237 bp in length
          *    9079     9178    gap of unknown length
          *    9179    10182    contig of 1004 bp in length
          *   10183    10282    gap of unknown length
          *   10283    11878    contig of 1596 bp in length
          *   11879    11978    gap of unknown length
          *   11979    13014    contig of 1036 bp in length
          *   13015    13114    gap of unknown length
          *   13115    14445    contig of 1331 bp in length
          *   14446    14545    gap of unknown length
          *   14546    15626    contig of 1081 bp in length
```

```
*  15627  15726: gap of unknown length
*  15727  17725: contig of 1999 bp in length
*  17726  17825: gap of unknown length
*  17826  18834: contig of 1009 bp in length
*  18835  18934: gap of unknown length
*  18935  20017: contig of 1083 bp in length
*  20018  20117: gap of unknown length
*  20118  21312: contig of 1195 bp in length
*  21313  21412: gap of unknown length
*  21413  22441: contig of 1029 bp in length
*  22442  22541: gap of unknown length
*  22542  24150: contig of 1609 bp in length
*  24151  24250: gap of unknown length
*  24251  25576: contig of 1326 bp in length
*  25577  25676: gap of unknown length
*  25677  27006: contig of 1330 bp in length
*  27007  27106: gap of unknown length
*  27107  28396: contig of 1290 bp in length
*  28397  28496: gap of unknown length
*  28497  30060: contig of 1564 bp in length
*  30061  30160: gap of unknown length
*  30161  31555: contig of 1395 bp in length
*  31556  31655: gap of unknown length
*  31656  34141: contig of 2486 bp in length
*  34142  34241: gap of unknown length
*  34242  36411: contig of 2170 bp in length
*  36412  36511: gap of unknown length
*  36512  38068: contig of 1557 bp in length
*  38069  38168: gap of unknown length
*  38169  39714: contig of 1546 bp in length
*  39715  39814: gap of unknown length
*  39815  41252: contig of 1438 bp in length
*  41253  41352: gap of unknown length
*  41353  42561: contig of 1209 bp in length
*  42562  42661: gap of unknown length
*  42662  44468: contig of 1807 bp in length
*  44469  44568: gap of unknown length
*  44569  46135: contig of 1567 bp in length
*  46136  46235: gap of unknown length
*  46236  47865: contig of 1630 bp in length
*  47866  47965: gap of unknown length
*  47966  49673: contig of 1708 bp in length
*  49674  49773: gap of unknown length
*  49774  51368: contig of 1595 bp in length
*  51369  51468: gap of unknown length
*  51469  52801: contig of 1333 bp in length
*  52802  52901: gap of unknown length
*  52902  54564: contig of 1663 bp in length
*  54565  54664: gap of unknown length
*  54665  57404: contig of 2740 bp in length
*  57405  57504: gap of unknown length
*  57505  59524: contig of 2020 bp in length
*  59525  59624: gap of unknown length
*  59625  62261: contig of 2637 bp in length
*  62262  62361: gap of unknown length
*  62362  64918: contig of 2557 bp in length
*  64919  65018: gap of unknown length
*  65019  68220: contig of 3202 bp in length
*  68221  68320: gap of unknown length
*  68321  70023: contig of 1703 bp in length
*  70024  70123: gap of unknown length
*  70124  72638: contig of 2515 bp in length
*  72639  72738: gap of unknown length
*  72739  74842: contig of 2104 bp in length
*  74843  74942: gap of unknown length
*  74943  76989: contig of 2047 bp in length
*  76990  77089: gap of unknown length
*  77090  79569: contig of 2480 bp in length
*  79570  79669: gap of unknown length
*  79670  81513: contig of 1844 bp in length
*  81514  81613: gap of unknown length
*  81614  84201: contig of 2588 bp in length
*  84202  84301: gap of unknown length
```

```
*  84302  87194: contig of 2893 bp in length
*  87195  87294: gap of unknown length
*  87295  89564: contig of 2270 bp in length
*  89565  89664: gap of unknown length
*  89665  92844: contig of 3180 bp in length
*  92845  92944: gap of unknown length
*  92945  95216: contig of 2272 bp in length
*  95217  95316: gap of unknown length
*  95317  98497: contig of 3181 bp in length
*  98498  98597: gap of unknown length
*  98598 101034: contig of 2437 bp in length
* 101035 101134: gap of unknown length
* 101135 104185: contig of 3051 bp in length
* 104186 104285: gap of unknown length
* 104286 107830: contig of 3545 bp in length
* 107831 107930: gap of unknown length
* 107931 111290: contig of 3360 bp in length
* 111291 111390: gap of unknown length
```

```
Query Match              10.5%;  Score 19;  DB 2;  Length 176036;
Best Local Similarity   100.0%;  Pred. No. 12;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       41 GAAGGATTTTATTACTTA 59
            ||||||||||||||||||
Db   104532 GAAGGATTTTATTACTTA 104550


RESULT 35
AC127774/c
LOCUS       AC127774              180153 bp    DNA     linear   HTG 19-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-89B16, *** SEQUENCING IN PROGRESS
            ***, 59 unordered pieces.
ACCESSION   AC127774
VERSION     AC127774.1  GI:21908146
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 180153)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,N.P., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,A.,
            Massey,J., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
```

```
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 180153)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     ------------------ Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ------------------ Project Information
            Center project name: GZSQ
            Center clone name: CH230-89B16
            ------------------ Summary Statistics
            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 133210 bases at least Q40
            Consensus quality: 140812 bases at least Q30
            Consensus quality: 146966 bases at least Q20
            ------------------
          * NOTE: Estimated insert size may differ from sequence length
          *    (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 59 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1      1420: contig of 1420 bp in length
          *    1421      1520: gap of unknown length
          *    1521      2931: contig of 1411 bp in length
          *    2932      3031: gap of unknown length
          *    3032      4421: contig of 1390 bp in length
          *    4422      4521: gap of unknown length
          *    4522      5902: contig of 1381 bp in length
          *    5903      6002: gap of unknown length
          *    6003      7241: contig of 1239 bp in length
          *    7242      7341: gap of unknown length
          *    7342      9500: contig of 2159 bp in length
          *    9501      9600: gap of unknown length
          *    9601     10646: contig of 1046 bp in length
          *   10647     10746: gap of unknown length
          *   10747     11830: contig of 1084 bp in length
          *   11831     11930: gap of unknown length
          *   11931     13178: contig of 1248 bp in length
          *   13179     13278: gap of unknown length
          *   13279     14730: contig of 1452 bp in length
          *   14731     14830: gap of unknown length
          *   14831     16464: contig of 1634 bp in length
          *   16465     16564: gap of unknown length
          *   16565     18209: contig of 1645 bp in length
          *   18210     18309: gap of unknown length
          *   18310     20250: contig of 1941 bp in length
          *   20251     20350: gap of unknown length
          *   20351     21750: contig of 1400 bp in length
          *   21751     21850: gap of unknown length
          *   21851     23381: contig of 1531 bp in length
          *   23382     23481: gap of unknown length
```

```
          *   23482     24653: contig of 1172 bp in length
          *   24654     24753: gap of unknown length
          *   24754     26173: contig of 1420 bp in length
          *   26174     26273: gap of unknown length
          *   26274     27958: contig of 1685 bp in length
          *   27959     28059: gap of unknown length
          *   28059     29402: contig of 1344 bp in length
          *   29403     29502: gap of unknown length
          *   29503     31187: contig of 1685 bp in length
          *   31188     31287: gap of unknown length
          *   31288     33042: contig of 1755 bp in length
          *   33043     33142: gap of unknown length
          *   33143     35128: contig of 1986 bp in length
          *   35129     35228: gap of unknown length
          *   35229     38085: contig of 2857 bp in length
          *   38086     38185: gap of unknown length
          *   38186     39781: contig of 1596 bp in length
          *   39782     39881: gap of unknown length
          *   39882     42266: contig of 2385 bp in length
          *   42267     42366: gap of unknown length
          *   42367     43566: contig of 1200 bp in length
          *   43567     43666: gap of unknown length
          *   43667     45411: contig of 1745 bp in length
          *   45412     45511: gap of unknown length
          *   45512     47033: contig of 1522 bp in length
          *   47034     47133: gap of unknown length
          *   47134     49387: contig of 2254 bp in length
          *   49388     49487: gap of unknown length
          *   49488     51080: contig of 1593 bp in length
          *   51081     51180: gap of unknown length
          *   51181     53455: contig of 2275 bp in length
          *   53456     53555: gap of unknown length
          *   53555     55559: contig of 2004 bp in length
          *   55560     55659: gap of unknown length
          *   55660     58007: contig of 2348 bp in length
          *   58008     58108: gap of unknown length
          *   58108     61502: contig of 3395 bp in length
          *   61503     61602: gap of unknown length
          *   61603     63437: contig of 1835 bp in length
          *   63438     63537: gap of unknown length
          *   63538     66288: contig of 2751 bp in length
          *   66288     66388: gap of unknown length
          *   66389     68546: contig of 2158 bp in length
          *   68547     68646: gap of unknown length
          *   68647     73537: contig of 4891 bp in length
          *   73538     73637: gap of unknown length
          *   73638     77125: contig of 3488 bp in length
          *   77126     77225: gap of unknown length
          *   77226     80563: contig of 3338 bp in length
          *   80564     80663: gap of unknown length
          *   80664     83984: contig of 3321 bp in length
          *   83985     84084: gap of unknown length
          *   84085     87419: contig of 3335 bp in length
          *   87420     87519: gap of unknown length
          *   87520     91100: contig of 3581 bp in length
          *   91101     91200: gap of unknown length
          *   91201     95131: contig of 3931 bp in length
          *   95132     95231: gap of unknown length
          *   95232     98354: contig of 3123 bp in length
          *   98355     98454: gap of unknown length
          *   98455    104119: contig of 5665 bp in length
          *  104120    104219: gap of unknown length
          *  104220    107789: contig of 3570 bp in length
          *  107790    107889: gap of unknown length
          *  107890    111584: contig of 3695 bp in length
          *  111585    111684: gap of unknown length
          *  111685    116567: contig of 4883 bp in length
          *  116568    116667: gap of unknown length
          *  116668    119994: contig of 3327 bp in length
          *  119995    120094: gap of unknown length
          *  120095    124984: contig of 4890 bp in length
          *  124985    125084: gap of unknown length
          *  125085    129333: contig of 4249 bp in length
```

```
      1    1567  *   1566: contig of 1566 bp in length
   1567    1667  *   1666: gap of unknown length
   1667    2773  *   2772: contig of 1106 bp in length
   2773    2873  *   2872: gap of unknown length
   2873    5021  *   5020: contig of 2148 bp in length
   5021    5121  *   5120: gap of unknown length
   5121    7359  *   7358: contig of 2238 bp in length
   7359    7459  *   7458: gap of unknown length
   7459    9056  *   9055: contig of 1597 bp in length
   9056    9156  *   9155: gap of unknown length
   9156   12154  *  12153: contig of 2998 bp in length
  12154   12254  *  12253: gap of unknown length
  12254   13643  *  13642: contig of 1389 bp in length
  13643   13743  *  13742: gap of unknown length
  13743   15773  *  15772: contig of 2030 bp in length
  15773   15873  *  15872: gap of unknown length
  15873   18137  *  18136: contig of 2264 bp in length
  18137   18237  *  18236: gap of unknown length
  18237   20008  *  20007: contig of 1771 bp in length
  20008   20108  *  20107: gap of unknown length
  20108   22203  *  22202: contig of 2095 bp in length
  22203   22303  *  22302: gap of unknown length
  22303   25428  *  25427: contig of 3125 bp in length
  25428   25528  *  25527: gap of unknown length
  25528   29039  *  29038: contig of 3511 bp in length
  29039   29139  *  29138: gap of unknown length
  29139   32006  *  32005: contig of 2867 bp in length
  32006   32106  *  32105: gap of unknown length
  32106   34858  *  34857: contig of 2752 bp in length
  34858   34958  *  34957: gap of unknown length
  34958   37235  *  37234: contig of 2277 bp in length
  37235   37335  *  37334: gap of unknown length
  37335   39702  *  39701: contig of 2367 bp in length
  39702   39802  *  39801: gap of unknown length
  39802   42194  *  42193: contig of 2392 bp in length
  42194   42294  *  42293: gap of unknown length
  42294   44857  *  44856: contig of 2563 bp in length
  44857   44957  *  44956: gap of unknown length
  44957   45486  *  45485: contig of 529 bp in length
  45486   45586  *  45585: gap of unknown length
  45586   48183  *  48182: contig of 2597 bp in length
  48183   48283  *  48282: gap of unknown length
  48283   51588  *  51587: contig of 3305 bp in length
  51588   51688  *  51687: gap of unknown length
  51688   55167  *  55166: contig of 3479 bp in length
  55167   55267  *  55266: gap of unknown length
  55267   58903  *  58902: contig of 3636 bp in length
  58903   59003  *  59002: gap of unknown length
  59003   63146  *  63145: contig of 4143 bp in length
  63146   63246  *  63245: gap of unknown length
  63246   67335  *  67334: contig of 4089 bp in length
  67335   67435  *  67434: gap of unknown length
  67435   71884  *  71883: contig of 4449 bp in length
  71884   71984  *  71983: gap of unknown length
  71984   75659  *  75658: contig of 3675 bp in length
  75659   75759  *  75758: gap of unknown length
  75759   80421  *  80420: contig of 4662 bp in length
  80421   80521  *  80520: gap of unknown length
  80521   84658  *  84657: contig of 4137 bp in length
  84658   84758  *  84757: gap of unknown length
  84758   89092  *  89091: contig of 4334 bp in length
  89092   89192  *  89191: gap of unknown length
  89192   94039  *  94038: contig of 4847 bp in length
  94039   94139  *  94138: gap of unknown length
  94139   99557  *  99556: contig of 5418 bp in length
  99557   99657  *  99656: gap of unknown length
  99657  105663  * 105662: contig of 6006 bp in length
 105663  105763  * 105762: gap of unknown length
 105763  112128  * 112127: contig of 6365 bp in length
 112128  112228  * 112227: gap of unknown length
 112228  118548  * 118547: contig of 6320 bp in length
 118548  118648  * 118647: gap of unknown length
 118648  125755  * 125755: contig of 7108 bp in length
```

```
 129334  129433  *  129433: gap of unknown length
 129434  133919  *  133919: contig of 4486 bp in length
 133920  134019  *  134019: gap of unknown length
 134020  140279  *  140279: contig of 6260 bp in length
 140280  140379  *  140379: gap of unknown length
 140380  145752  *  145752: contig of 5373 bp in length
 145753  145852  *  145852: gap of unknown length
 145853  151624  *  151624: contig of 5772 bp in length
 151625  151724  *  151724: gap of unknown length
 151725  158653  *  158653: contig of 6929 bp in length
 158654  158753  *  158753: gap of unknown length
```

Query Match   10.5%;   Score 19;   DB 2;   Length 180153;
Best Local Similarity   100.0%;   Pred. No. 12;
Matches 19;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      41 GAAGGATTTTTATTACTTA 59
           ||||||||||||||||||||
Db   66091 GAAGGATTTTTTATTACTTA 66073
```

RESULT 36
AC024134
LOCUS   AC024134   183178 bp   DNA   linear   HTG 01-SEP-2000
DEFINITION Homo sapiens chromosome 17 clone RP11-68I05, *** SEQUENCING IN
            PROGRESS ***, 48 unordered pieces.
ACCESSION   AC024134
VERSION   AC024134.4   GI:9958316
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 183178)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 183178)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (24-FEB-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT   On Sep 1, 2000 this sequence version replaced gi:8844178.

            -------------- Genome Center --------------
            Center: Washington University Gnome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0681O05
            -------------- Summary Statistics --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 155289 bases at least Q40
            Consensus quality: 162368 bases at least Q30
            Consensus quality: 166518 bases at least Q20
            Insert size: 172000; agarose-fp
            Insert size: 177949; sum-of-contigs
            Quality coverage: 2.85 in Q20 bases; agarose-fp
            Quality coverage: 3.19 in Q20 bases; sum-of-contigs

          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 48 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.

```
     *      125756            125855: gap of unknown length
     *      125856            130812: contig of 4957 bp in length
     *      130813            130912: gap of unknown length
     *      130913            138250: contig of 7338 bp in length
     *      138251            138350: gap of unknown length
     *      138351            148729: contig of 10379 bp in length
     *      148730            148829: gap of unknown length
     *      148830            158131: contig of 9302 bp in length
     *      158132            158231: gap of unknown length
     *      158232            173848: contig of 15617 bp in length
     *      173849            173948: gap of unknown length
     *      173949            175348: contig of 1400 bp in length
     *      175349            175448: gap of unknown length
     *      175449            176535: contig of 1087 bp in length
     *      176536            176635: gap of unknown length
     *      176636            178205: contig of 1570 bp in length
     *      178206            178305: gap of unknown length
     *      178306            179753: contig of 1448 bp in length
     *      179754            179853: gap of unknown length
     *      179854            181517: contig of 1664 bp in length
     *      181518            181617: gap of unknown length
     *      181618            183178: contig of 1561 bp in length.
FEATURES             Location/Qualifiers
     source          1..183178
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="RP11-68105"
     misc_feature    1..1566
                     /note="assembly_name:Contig100"
     misc_feature    1667..2772
                     /note="assembly_name:Contig102"
     misc_feature    2873..5020
                     /note="assembly_name:Contig103"
     misc_feature    5121..7358
                     /note="assembly_name:Contig104"
     misc_feature    7459..9055
                     /note="assembly_name:Contig105"
     misc_feature    9156..12153
                     /note="assembly_name:Contig106"
     misc_feature    12254..13642
                     /note="assembly_name:Contig107"
     misc_feature    13743..15772
                     /note="assembly_name:Contig108"
     misc_feature    15873..18136
                     /note="assembly_name:Contig109"
     misc_feature    18237..20007
                     /note="assembly_name:Contig110"
     misc_feature    20108..22202
                     /note="assembly_name:Contig112"
     misc_feature    22303..25427
                     /note="assembly_name:Contig113"
     misc_feature    25528..29038
                     /note="assembly_name:Contig114"
     misc_feature    29139..32005
                     /note="assembly_name:Contig121
                     clone_end:SP6
                     vector_side:right"
     misc_feature    32106..34857
                     /note="assembly_name:Contig116"
     misc_feature    34958..37234
                     /note="assembly_name:Contig117"
     misc_feature    37335..39701
                     /note="assembly_name:Contig118"
     misc_feature    39802..42193
                     /note="assembly_name:Contig119"
     misc_feature    42294..44856
                     /note="assembly_name:Contig120"
     misc_feature    44957..45485
                     /note="assembly_name:Contig53"
     misc_feature    45586..48182
                     /note="assembly_name:Contig122"
     misc_feature    48283..51587
```

```
     misc_feature    51688..55166
                     /note="assembly_name:Contig123"
     misc_feature    55267..58902
                     /note="assembly_name:Contig124"

Query Match             10.5%;    Score 19;  DB 2;  Length 183178;
Best Local Similarity  100.0%;    Pred. No. 12;
Matches   19; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy        111 CAAAGTTGATGCTGAGAAC 129
              ||||||||||||||||||
Db      24459 CAAAGTTGATGCTGAGAAC 24477


RESULT 37
AC073162
LOCUS       AC073162          185754 bp    DNA     linear   PRI 26-APR-2002
DEFINITION  Homo sapiens chromosome 10 clone RP11-536J24, complete sequence.
ACCESSION   AC073162
VERSION     AC073162.9  GI:20330787
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 185754)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185754)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   3  (bases 1 to 185754)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-APR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   4  (bases 1 to 185754)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REMARK      1329 bases of Tn10 (J01829) transposon removed here
COMMENT     On Apr 26, 2002 this sequence version replaced gi:20066186.
FEATURES             Location/Qualifiers
     source          1..185754
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-536J24"
                     /clone_lib="RPCI-11"
BASE COUNT    61914 a    38909 c    36605 g    48326 t
ORIGIN

Query Match             10.5%;    Score 19;  DB 9;  Length 185754;
Best Local Similarity  100.0%;    Pred. No. 12;
Matches   19; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy         49 TTTATTACTTATATGGACA 67
              ||||||||||||||||||
Db      13249 TTTATTACTTATATGGACA 13267


RESULT 38
AC006281/c
LOCUS       AC006281          199551 bp    DNA     linear   HTG 12-AUG-2000
DEFINITION  Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
            PROGRESS ***, 2 unordered pieces.
ACCESSION   AC006281
VERSION     AC006281.8  GI:9797738
```

```
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Plasmodium falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 199551)
  AUTHORS   Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
            Kurdi,O.B., Conway,A.B. and Davin,R.W.
  TITLE     Plasmodium falciparum 3D7 chromosome 12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199551)
  AUTHORS   Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi.O.B. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
            Center, Stanford University, 855 California Avenue, Palo Alto, CA
            94304, USA
COMMENT     On Aug 12, 2000 this sequence version replaced gi:8810457.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1       41526: contig of 41526 bp in length
            *    41527       41726: gap of unknown length
            *    41727      199551: contig of 157825 bp in length.
FEATURES             Location/Qualifiers
     source          1..199551
                     /organism="Plasmodium falciparum"
                     /db_xref="taxon:5833"
                     /chromosome="12"
                     /clone="PFYACB8-95"
                     /clone="3D7"
BASE COUNT    83197 a   17543 c   19759 g   78852 t    200 others
ORIGIN

Query Match        10.5%;   Score 19;   DB 2;   Length 199551;
Best Local Similarity  100.0%;   Pred. No. 12;
Matches   19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        46 ATTTTTATTACTTATATGG 64
             |||||||||||||||||||
Db     66147 ATTTTTATTACTTATATGG 66129
```

```
RESULT 39
HUAE000659
LOCUS       HUAE000659            251199 bp    DNA     linear   PRI 15-JUN-2001
DEFINITION  Homo sapiens T-cell receptor alpha delta locus from bases 250472 to
            501670 (section 2 of 5) of the Complete Nucleotide Sequence.
ACCESSION   AE000659 AE000521 U85199
VERSION     AE000659.1  GI:2358025
KEYWORDS    .
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 251199)
  AUTHORS   Koop,B.F., Rowen,L., Wang,K., Kuo,C.L., Seto,D., Lenstra,J.A.,
            Howard,S., Shan,W., Deshpande,P. and Hood,L.
  TITLE     The human T-cell receptor TCRAC/TCRDC (C alpha/C delta) region:
            organization, sequence, and evolution of 97.6 kb of DNA
  JOURNAL   Genomics 19 (3), 478-493 (1994)
  MEDLINE   94245236
  PUBMED    8188290
  REMARK    This citation covers from bases 966383-1064019
REFERENCE   2  (bases 1 to 251199)
  AUTHORS   Boysen,C., Simon,M.I. and Hood,L.
  TITLE     Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with
            bacterial artificial chromosome clones
  JOURNAL   Genome Res. 7 (4), 330-338 (1997)
  MEDLINE   97264339
  PUBMED    9110172
REFERENCE   3  (bases 1 to 251199)
  AUTHORS   Boysen,C., Inyoul,L., Smith,T.M., Smit,A., Wang,K., Rowen,L. and
            Hood,L.
  TITLE     T-Cell Receptor Alpha Delta Locus Complete Nucleotide Sequence
  JOURNAL   Unpublished
  REMARK    This citation covers bases 1-983545 and bases 1064020-1071650
REFERENCE   4  (bases 1 to 251199)
  AUTHORS   Boysen,C., Lee,I., Smith,T.M., Smit,A., Kai,W., Lee,R. and Leroy,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-1997) Department of Molecular Biotechnology,
            University of Washington, Box 357730, Seattle, Washington 98195,
            USA
  REMARK    Complete nucleotide sequence of the human T-cell receptor alpha
            delta locus
COMMENT     On Jun 12, 2001 this sequence version replaced gi:2766593.
            Sequencing method: high redundancy shotgun. Interspersed Repeats
            were identified with RepeatMasker (available from
            http://ftp.genome.washington.edu/RM/RepeatMasker.html) Simple
            sequence repeats were identified with sputnik (available from
            http://serac.mbt.washington.edu/~chrisa/software/sputnik.html).
FEATURES             Location/Qualifiers
     source          1..251199
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     source          33706..138937
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Mel Simon's BAC library"
                     /clone="BAC 956"
     repeat_region   458..1229
                     /rpt_family="L1"
     repeat_region   1252..1542
                     /rpt_family="AluSp"
     repeat_region   1711..1791
                     /rpt_family="MER5A"
     repeat_region   2008..2207
                     /rpt_family="MER20"
     gene            4449..5201
                     /gene="TCRAV5S1"
     promoter        4449..4468
                     /gene="TCRAV5S1"
     mRNA            join(<4647..4698,4875..>5162)
                     /gene="TCRAV5S1"
                     /product="TCRAV5S1"
                     /note="Old name TCRAV15S1"
     CDS             join(4647..4698,4875..>5162)
                     /gene="TCRAV5S1"
                     /standard_name="TCRAV5S1"
                     /note="Old name TCRAV15S1"
                     /codon_start=1
                     /product="TCRAV5S1"
                     /protein_id="AAB69002.1"
                     /db_xref="GI:2358026"
                     /translation="MKTPAGFSFLFLWLQLDCMSRGEDVEQSLFLSVREGDSSVINCT
                     YTDSSSTYLYWYKQEPGAGLQLLTYIFSNMDMKQDQRLTVLLNKKDKHLSLRIADTQT
                     GDSAIYPCAES"
     v_segment       join(4647..4698,4875..>5162)
                     /gene="TCRAV5S1"
                     /standard_name="TCRAV5S1"
                     /note="Old name TCRAV15S1"
     misc_recomb     5163..5169
                     /gene="TCRAV5S1"
                     /note="heptamer"
     misc_recomb     5170..5192
                     /gene="TCRAV5S1"
                     /standard_name="TCRAV5S1"
                     /note="spacer"
     misc_recomb     5193..5201
                     /gene="TCRAV5S1"
```

```
     repeat_region    5650..5724
                      /standard_name="TCRAV5S1"
                      /note="nonamer"
     repeat_region    5650..5724
                      /rpt_family="MER4_internal"
     repeat_region    complement(7864..8152)
                      /rpt_family="AluY"
     repeat_region    8253..8262
                      /rpt_type=tandem
                      /rpt_unit=ATC
     repeat_region    8886..8892
                      /rpt_type=tandem
                      /rpt_unit=TTTA
     repeat_region    complement(8895..9195)
                      /rpt_family="AluSq"
     repeat_region    9236..10947
                      /rpt_family="L1P4_5end"
     repeat_region    10101..11383
                      /rpt_family="L1"
     repeat_region    complement(11565..11901)
                      /rpt_family="L1PA4"
     repeat_region    complement(11892..12127)
                      /rpt_family="L1PA3"
     repeat_region    complement(11978..15338)
                      /rpt_family="L1"
     repeat_region    15339..15381
                      /rpt_family="AluJb"
     repeat_region    15531..20950
                      /rpt_family="L1"
     repeat_region    20801..21692
                      /rpt_family="L1PA2"
     repeat_region    complement(22489..22718)
                      /rpt_family="MIR"
     repeat_region    23528..23680
                      /rpt_family="MER5A"
     gene             23719..24532
                      /gene="TCRAV6S1"
     promoter         23719..23738
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="Old name TCRAV5S1"
     mRNA             join(<23966..24017,24206..>24493)
                      /gene="TCRAV6S1"
                      /product="TCRAV6S1"
                      /note="Old name TCRAV6S1"
     CDS              join(23966..24017,24206..>24493)
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="Old name TCRAV6S1"
                      /codon_start=1
                      /product="TCRAV6S1"
                      /protein_id="AAB69003.1"
                      /db_xref="GI:2358027"
                      /translation="MESFLGGVLLILMLQVDWVKSQKIEQNSEALNIQEGKTATLTCN
                      YTNYSPAYLQWYRQDPGRGPVFLLIRENEKEKRKERLKVTFDTTLKQSLFHITASQP
                      ADSATYLCALD"
     V_segment        join(23966..24017,24206..>24493)
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="Old name TCRAV6S1"
     misc_recomb      24494..24500
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="heptamer"
     misc_recomb      24501..24523
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="spacer"
     misc_recomb      24524..24532
                      /gene="TCRAV6S1"
                      /standard_name="TCRAV6S1"
                      /note="nonamer"
     repeat_region    complement(25031..25204)
                      /rpt_family="AluJo"

     repeat_region    26126..26302
                      /rpt_family="MIR"
     repeat_region    26503..26734
                      /rpt_family="MIR2"
     repeat_region    26841..27388
                      /rpt_family="LTR8"
     repeat_region    complement(28325..28702)
                      /rpt_family="MIR2"
     repeat_region    30840..31338
                      /rpt_family="L1"
     repeat_region    31187..31459
                      /rpt_family="L1ME2"
     repeat_region    31927..31981
                      /rpt_family="MIR2"
     repeat_region    32185..32727
                      /rpt_family="MER41A"
     repeat_region    complement(32845..33150)
                      /rpt_family="AluY"
     repeat_region    complement(33208..33371)
                      /rpt_family="L1MC1"
     repeat_region    complement(33812..34156)
                      /rpt_family="L1_5end"
     repeat_region    complement(34365..34657)
                      /rpt_family="AluSx"
     repeat_region    35255..35570
                      /rpt_family="AluSx"
     repeat_region    complement(35929..36158)
                      /rpt_family="MER52"
     repeat_region    complement(36067..36235)


Query Match     10.5%;  Score 19;  DB 9;  Length 251199;
Best Local Similarity   100.0%;  Pred. No. 11;
Matches  19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       42 AAGGATTTTTATTACTTAT 60
            |||||||||||||||||||
Db   216090 AAGGATTTTTATTACTTAT 216108


RESULT 40
BRA417800/c
LOCUS       BRA417800              410 bp    DNA     linear   PLN 31-MAY-2002
DEFINITION  Brassica rapa subsp. oleifera DNA for reverse transcriptase gene of
            Line retroelements, 12L41-07.
ACCESSION   AJ417800
VERSION     AJ417800.1  GI:18564974
KEYWORDS    LINE element; reverse transcriptase.
SOURCE      Brassica rapa subsp. oleifera
ORGANISM    Brassica rapa subsp. oleifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Alix,K., Ryder,C., King,G. and Heslop-Harrison,J.S.
  TITLE     The genomic organization of retroelements in Brassica oleracea
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Alix,K. and Heslop-Harrison,J.S.
  TITLE     The Diversity of Retroelements in Diploid and Tetraploid Brassica
            Species
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 410)
  AUTHORS   Alix,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2001) Alix K., Biology, University of Leicester,
            University Road, LE1 7RH, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1..410
                     /organism="Brassica rapa subsp. oleifera"
                     /cultivar="R 500"
                     /sub_species="oleifera"
                     /db_xref="taxon:145471"
                     /note="id 12L41-07"
```

```
CDS              <1..>410
                 /codon_start=1
                 /product="reverse transcriptase"
                 /protein_id="CAD22562.1"
                 /db_xref="GI:18564975"
                 /db_xref="SPTREMBL:Q8VXT5"
                 /translation="NDFRPISCCHIVYKVISKIIANRLKPILRDCVSPNQAAFLKGRS
                 LGENVLASELIRDYNKSSCLRSSMLKVDIRKAPDTVCWDFVIKILRAQGFPPLFVSW
                 IQECITSPRFSVALNGELAGFPFGKKKGLRQGDPLS"

BASE COUNT       95 a      90 c      99 g     126 t
ORIGIN

Query Match            9.9%;   Score 18;  DB 8;  Length 410;
Best Local Similarity  100.0%; Pred. No. 50;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      39 CCGAAGGATTTTTATTAC 56
           ||||||||||||||||||
Db     273 CCGAAGGATTTTTATTAC 256


RESULT 41
G39242/c
LOCUS       G39242              659 bp    DNA     linear   STS 01-FEB-2001
DEFINITION  Z20324 Zebrafish AB Danio rerio STS genomic, sequence tagged site.
ACCESSION   G39242
VERSION     G39242.1  GI:3358451
KEYWORDS    STS.
SOURCE      Danio rerio.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Shimoda,N., Knapik,E.W., Ziniti,J., Sim,C., Yamada,E., Kaplan,S.,
            Jackson,D., de Sauvage,F., Jacob,H. and Fishman,M.C.
  TITLE     Zebrafish genetic map with 2000 microsatellite markers
  JOURNAL   Genomics 58 (3), 219-232 (1999)
  MEDLINE   99303552
   PUBMED   10373319
COMMENT     Contact: Mark C. Fishman
            Cardiovascular Research Center
            Massachusetts General Hospital
            Mail code 1491100A, 149 13th Street, Charlestown, MA 02129, USA
            Fax: 6172658060
            Email: fishman@mgh.cvrc.harvard.edu
            http://zebrafish.mgh.harvard.edu
            Primer A: GTTGGGTGAAACAGTTGGG
            Primer B: TGGAAGGATCATCTGAACCC
            STS size: 138
            PCR Profile:
            Template:         10 ng
            Primer:           each 375 nM
            Denaturation:     94 degrees C for 5.0 minutes
            Denaturation:     94 degrees C for 1.0 minute
            Annealing:        58 degrees C for 1.0 minute
            Polymerization:   72 degrees C for 1.5 minute
            PCR Cycles:       27
            Thermal Cycler:   MJ Research PTC-100
            Protocol:
            dNTPs:            each 200 uM
            Taq Polymerase:   0.034 units/ul
            Total Vol:        10 ul

            Buffer:
                              MgCL2:     1.5 mM
                              KCl:       50 mM
                              Tris-HCl:  10 mM
                              pH:        8.3

            Primers are available from Research Genetics Inc.
            (http://www.resgen.com phone: 800-533-4363).
```

```
FEATURES             Location/Qualifiers
     source          1..659
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /sex="F"
                     /clone_lib="Zebrafish AB"
                     /dev_stage="Adult"
                     /lab_host="DH5alphaF'IQ"
                     /note="Vector: m13MP19 with added BstXI site; V-type:
                     Phage; Genomic DNA from a single adult Zebrafish of AB
                     strain was digested with AluI, Cac8I, HaeIII, NlaIV, or
                     RsaI. Fragments in the range of 250-500 bp were gel
                     purified and a BstXI linker was added. The fragments were
                     cloned into a modified M13mp19 vector and transformed
                     into E. Coli DH5alpha. Microsatellite sequences were
                     screened with labeled d(CA)15 and d(GT)15 oligonucleotide
                     probes."
     STS             137..274
     primer_bind     137..156
     primer_bind     complement(255..274)
BASE COUNT    124 a     117 c     218 g     183 t     17 others
ORIGIN

Query Match            9.9%;   Score 18;  DB 11;  Length 659;
Best Local Similarity  100.0%; Pred. No. 49;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      74 GCCAGAACAACAACAGCTG 91
           |||||||||||||||||||
Db      88 GCCAGAACAACAACAGCTG 71


RESULT 42
HSUATM32
LOCUS       HSUATM32              750 bp    DNA     linear   PRI 10-DEC-1996
DEFINITION  Human ataxia-telangiectasia (ATM) gene, exon 36.
ACCESSION   U55733 U40894
VERSION     U55733.1  GI:1497905
KEYWORDS    .
SEGMENT     32 of 56
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   Rasio,D., Negrini,M. and Croce,C.M.
  TITLE     Genomic organization of the ATM locus involved in
            ataxia-telangiectasia
  JOURNAL   Cancer Res. 55 (24), 6053-6057 (1995)
  MEDLINE   96105020
   PUBMED   8521392
REFERENCE   2  (bases 1 to 750)
  AUTHORS   Vorechovsky,I., Rasio,D., Luo,L., Monaco,C., Hammarstrom,L.,
            Webster,A.D.B., Zaloudik,J., Barbanti-Brodano,G., James,M.,
            Russo,G., Croce,C.M. and Negrini,M.
  TITLE     The ATM gene and susceptibility to breast cancer: analysis of 38
            breast tumors reveals no evidence for mutation
  JOURNAL   Cancer Res. 56 (12), 2726-2732 (1996)
  MEDLINE   96275738
   PUBMED   8665503
REFERENCE   3  (bases 1 to 750)
  AUTHORS   Negrini,M., Rasio,D. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1996) Massimo Negrini, Kimmel Cancer Institute,
            Thomas Jefferson University, 233 S 10th Street, Philadelphia, PA
            19107, USA
COMMENT     On Aug 21, 1996 this sequence version replaced gi:1185484.
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
```

```
intron               /map="11q23"
                     <1..350
                     /gene="ATM"
                     /number=35
exon                 351..492
                     /gene="ATM"
                     /number=36
intron               493..>750
                     /gene="ATM"
                     /number=36
BASE COUNT    245 a     93 c    140 g    272 t
ORIGIN

Query Match             9.9%;  Score 18;  DB 9;  Length 750;
Best Local Similarity  100.0%;  Pred. No. 49;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         46 ATTTTTATTACTTATATG 63
              ||||||||||||||||||
Db         42 ATTTTTATTACTTATATG 59

RESULT 43
AX414620
LOCUS       AX414620                 996 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 1611 from Patent WO0228891.
ACCESSION   AX414620
VERSION     AX414620.1  GI:21447077
KEYWORDS    .
SOURCE      Listeria monocytogenes ATCC 19115.
  ORGANISM  Listeria monocytogenes ATCC 19115
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 1611 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1..996
                     /organism="Listeria monocytogenes ATCC 19115"
                     /db_xref="taxon:176281"
BASE COUNT    211 a    259 c    225 g    301 t
ORIGIN

Query Match             9.9%;  Score 18;  DB 6;  Length 996;
Best Local Similarity  100.0%;  Pred. No. 49;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         41 GAAGGATTTTTATTACTT 58
              ||||||||||||||||||
Db        296 GAAGGATTTTTATTACTT 313

RESULT 44
PGHAGCG/c
LOCUS       PGHAGCG                 1841 bp    DNA     linear   BCT 15-JAN-1997
DEFINITION  P.gingivalis hagC gene encoding HagC hemagglutinin protein.
ACCESSION   Z27394
VERSION     Z27394.1  GI:510187
KEYWORDS    hagC gene; hemagglutinin protein.
SOURCE      Porphyromonas gingivalis
  ORGANISM  Porphyromonas gingivalis
            Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
            Porphyromonadaceae; Porphyromonas.
REFERENCE   1  (bases 1 to 1841)
  AUTHORS   Lepine,G. and Progulske-Fox,A.
  TITLE     Duplication and differential expression of hemagglutinin genes in
            Porphyromonas gingivalis
  JOURNAL   Oral Microbiol. Immunol. 11 (2), 65-78 (1996)
  MEDLINE   97096936
   PUBMED   8941757
REFERENCE   2  (bases 1 to 1841)
  AUTHORS   Lepine,G.
  JOURNAL   Thesis (1993) Oral Biology, University of Florida
REFERENCE   3 (bases 1 to 1841)
  AUTHORS   Lepine,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1993) Guylaine Lepine, Oral Biology, University
            of Florida, 1600 SW, Archer Road, Gainesville, Florida, 32610-0424,
            USA
FEATURES             Location/Qualifiers
     source          1..1841
                     /organism="Porphyromonas gingivalis"
                     /strain="381"
                     /db_xref="taxon:837"
                     /clone="GL7"
                     /clone_lib="Hind III digestion of 381 chromosomal DNA in
                     pUC18"
     -35_signal      314..319
     -10_signal      339..344
     gene            374..1426
                     /gene="hagC"
     CDS             374..1426
                     /gene="hagC"
                     /function="Might be involved in the colonization process
                     of P. gingivalis in the oral cavity."
                     /note="HagC hemagglutinin protein is 98.6% homologous to
                     hagB protein encoded by hagB gene also isolated from P.
                     gingivalis 381"
                     /codon_start=1
                     /transl_table=11
                     /product="HagC hemagglutinin protein"
                     /protein_id="CAA81786.1"
                     /db_xref="GI:510188"
                     /db_xref="SPTREMBL:Q51830"
                     /translation="MTAEIFSFSRLQNLEHYRFAKNVLTLCRTANIAKLNPKLPELEK
                     AIEMEDLALNPPVANELTPQVIALDEERDRAYQALMSRVRSYAFDEDSQLRNAARIE
                     DVAARYGNVIRMNYDKETAAIENFLTDLKGENIRPLVTKLGGVTALVDRLEKNNKAFAD
                     PFLRRLSTDQRGKYDVKALRAETDRTLVAVVRRMDSIDDMEPSPEIRALIELYNRLVA
                     NRRALLARRASYGEAAVEKRRAEIAEMLRPLLARIVEEKTAVFAGRTLGTGKNRHYL
                     ITFVAENGDEDRWYRINGEQLVYVPEDELPKPKPKKKKPASSTDTPSEPPVLPDPSQG
                     GSSSGGEQGSTGGGL"
     stem_loop       1488..1528
                     /note="Presence of two inverted repeats of 20 bp
                     (1488-1507 and 1509-1528)"
                     /function="Putative transcription-termination signal"
BASE COUNT    464 a    507 c    471 g    399 t
ORIGIN

Query Match             9.9%;  Score 18;  DB 1;  Length 1841;
Best Local Similarity  100.0%;  Pred. No. 48;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         24 GCTGGAAGACGTCTTCCG 41
              ||||||||||||||||||
Db       1627 GCTGGAAGACGTCTTCCG 1610

RESULT 45
AR049926/c
LOCUS       AR049926                1841 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 5 from patent US 5824791.
ACCESSION   AR049926
VERSION     AR049926.1  GI:5971918
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1841)
  AUTHORS   Progulske-Fox,A., Tumwasorn,S., Lepine,G., Han,N., Lantz,M. and
            Patti,J.M.
  TITLE     Cloned porphyromonas gingivalis genes and probes for the detection
            of periodontal disease
  JOURNAL   Patent: US 5824791-A 5 20-OCT-1998;
FEATURES             Location/Qualifiers
     source          1..1841
```

BASE COUNT   464 a   507 c   471 g   399 t
ORIGIN   /organism="unknown"

Query Match   9.9%;   Score 18;   DB 6;   Length 1841;
Best Local Similarity 100.0%;   Pred. No. 48;
Matches 18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    24  GCTGGAAGAGACGTCTTCCG 41
          ||||||||||||||||||||
Db  1627  GCTGGAAGAGACGTCTTCCG 1610
```

Search completed: June 9, 2003, 12:33:20
Job time : 1271 secs

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:50 ; Search time 474 Seconds
                                        (without alignments)
                                        1289.365 Million cell updates/sec

Title: US-09-550-163-1_COPY_80_100
Perfect score: 21
Sequence: 1 acttatccaattcacacag 21

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
       1: gb_ba:*
       2: gb_htg:*
       3: gb_in:*
       4: gb_om:*
       5: gb_ov:*
       6: gb_pat:*
       7: gb_ph:*
       8: gb_pl:*
       9: gb_pr:*
      10: gb_ro:*
      11: gb_sts:*
      12: gb_sy:*
      13: gb_un:*
      14: gb_vi:*
      15: em_ba:*
      16: em_fun:*
      17: em_hum:*
      18: em_in:*
      19: em_mu:*
      20: em_om:*
      21: em_or:*
      22: em_ov:*
      23: em_pat:*
      24: em_ph:*
      25: em_pl:*
      26: em_ro:*
      27: em_sts:*
      28: em_un:*
      29: em_vi:*
      30: em_htg_hum:*
      31: em_htg_inv:*
      32: em_htg_other:*
      33: em_htg_mus:*
      34: em_htg_pln:*
      35: em_htg_rod:*
      36: em_htg_mam:*
      37: em_htg_vrt:*
      38: em_sy:*
      39: em_htgo_hum:*
      40: em_htgo_mus:*
      41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	732	6	AX406939	Sequence
2	21	100.0	732	6	AX406941	Sequence
3	21	100.0	732	6	AX406943	Sequence
4	21	100.0	732	6	AX406945	Sequence
5	21	100.0	732	9	AF071002	Homo sapi
6	21	100.0	809	9	AF302095	Homo sapi
7	21	100.0	24608	9	AP000320	Homo sapi
8	21	100.0	100000	9	AP000052	Homo sapi
9	21	100.0	100000	9	AP000167	Homo sapi
10	21	100.0	100000	17	AP000120	Homo sapi
11	21	100.0	340000	9	AP001719	Homo sapi
12 c	20	95.2	99887	2	AC026481	Homo sapi
13	19.4	92.4	732	6	AX406947	Sequence
14	19.4	92.4	164456	2	AC078804	Homo sapi
15 c	18.4	87.6	64677	2	AC079102	Homo sapi
16 c	18.4	87.6	132444	9	AC091878	Homo sapi
17 c	18.4	87.6	138118	2	AC102005	Mus muscu
18	18.4	87.6	152010	9	AC026785	Homo sapi
19	18.4	87.6	157723	9	AC120989	Oryza sat
20	18.4	87.6	177108	2	AC026504	Homo sapi
21	18.4	87.6	180277	2	AC104270	Oryza sat
22	18.4	87.6	193905	2	AC124440	Mus muscu
23	18.4	87.6	218617	2	AL732521	Mus muscu
24 c	18.4	87.6	222806	9	AC018977	Homo sapi
25	18.4	87.6	308013	2	AC125047	Mus muscu
26	18.4	87.6	318539	2	AL772310	Mus muscu
27	17.8	84.8	37049	9	AC017026	Homo sapi
28	17.8	84.8	37754	3	U23519	Caenorhabdi
29	17.8	84.8	69156	2	AC090542	Homo sapi
30 c	17.8	84.8	69172	2	AC129960	Bos tauru
31 c	17.8	84.8	89642	2	AC005136	Homo sapi
32	17.8	84.8	89820	9	AC007868	Genomic S
33 c	17.8	84.8	94991	2	AC097244	Rattus no
34	17.8	84.8	95597	9	AC012555	Homo sapi
35	17.8	84.8	107627	9	AL513218	Human DNA
36 c	17.8	84.8	110000	2	AC009771_1	Continuation (2 of
37 c	17.8	84.8	110000	2	AC012386_2	Continuation (3 of
38	17.8	84.8	113693	9	HSDJC16	Human DNA
39 c	17.8	84.8	115863	9	HS268D13	Human DNA
40 c	17.8	84.8	115915	9	AL023513	Human DNA
41	17.8	84.8	124048	9	AC105922	Human Chr
42 c	17.8	84.8	125419	2	HUAC002045	Human DNA
43	17.8	84.8	128294	9	AC026260	Homo sapi
44	17.8	84.8	128461	9	AC073248	Homo sapi
45 c	17.8	84.8	134443	2	AL390741	Human DNA
					AC015547	Homo sapi

ALIGNMENTS

RESULT 1
AX406939
LOCUS   AX406939   732 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION   Sequence 1 from Patent WO0222875.
ACCESSION   AX406939
VERSION   AX406939.1 GI:21439814
KEYWORDS   .
SOURCE   human.
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Goldstein,S.A.
  TITLE   Polymorphisms associated with cardiac arrythmia
  JOURNAL   Patent: WO 0222875-A 1 21-MAR-2002;

```
FEATURES             Location/Qualifiers
     source          YALE UNIVERSITY (US)
                     1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="minK-related peptide 1, Val substituted for Ala at
                     amino acid 116"
                     /codon_start=1
                     /protein_id="CAD35310.1"
                     /db_xref="GI:21439815"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGVAGFKMSP"
     variation       420
                     /note="The drug associated here was quinidine."
                     /replace="c"
BASE COUNT        221 a    151 c    157 g    203 t
ORIGIN
Qy          1 ACTTTATCCAATTCCACACAG 21
              |||||||||||||||||||||
Db         80 ACTTTATCCAATTCCACACAG 100

Query Match         100.0%;  Score  21;  DB  6;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 2
AX406941
LOCUS       AX406941             732 bp    DNA      linear   PAT 14-JUN-2002
DEFINITION  Sequence 3 from Patent WO0222875.
ACCESSION   AX406941
VERSION     AX406941.1  GI:21439816
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 3 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="minK-related peptide 1, wild type"
                     /codon_start=1
                     /protein_id="CAD35180.1"
                     /db_xref="GI:21439817"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
BASE COUNT        221 a    152 c    157 g    202 t
ORIGIN
Qy          1 ACTTTATCCAATTCCACACAG 21
              |||||||||||||||||||||
Db         80 ACTTTATCCAATTCCACACAG 100

Query Match         100.0%;  Score  21;  DB  6;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 3
AX406943
LOCUS       AX406943             732 bp    DNA      linear   PAT 14-JUN-2002
DEFINITION  Sequence 5 from Patent WO0222875.
ACCESSION   AX406943
VERSION     AX406943.1  GI:21439818
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 5 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="minK-related peptide 1, Thr substitued for Met at
                     amino acid 54"
                     /codon_start=1
                     /protein_id="CAD35181.1"
                     /db_xref="GI:21439819"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLTVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
     variation       234
                     /note="The drug associated here was procainamide."
                     /replace="t"
BASE COUNT        221 a    153 c    157 g    201 t
ORIGIN
Qy          1 ACTTTATCCAATTCCACACAG 21
              |||||||||||||||||||||
Db         80 ACTTTATCCAATTCCACACAG 100

Query Match         100.0%;  Score  21;  DB  6;  Length 732;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 4
AX406945
LOCUS       AX406945             732 bp    DNA      linear   PAT 14-JUN-2002
DEFINITION  Sequence 7 from Patent WO0222875.
ACCESSION   AX406945
VERSION     AX406945.1  GI:21439820
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Goldstein,S.A.
TITLE       Polymorphisms associated with cardiac arrythmia
JOURNAL     Patent: WO 0222875-A 7 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="minK-related peptide 1, Thr substituted for Ile at
                     amino acid 57"
                     /codon_start=1
                     /protein_id="CAD35311.1"
                     /db_xref="GI:21439821"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVWTGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"
     variation       243
                     /note="The drug associated here was oxatomide."
                     /replace="t"
BASE COUNT        221 a    153 c    157 g    201 t
```

```
ORIGIN

Query Match      100.0%;   Score 21; DB 6;   Length 732;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 ACTTTATCCAATTCACACAG 21
          ||||||||||||||||||||
Db     80 ACTTTATCCAATTCACACAG 100

RESULT 5
AF071002
LOCUS       AF071002                 732 bp    mRNA    linear   PRI 29-APR-1999
DEFINITION  Homo sapiens minK-related peptide 1 mRNA, complete cds.
ACCESSION   AF071002
VERSION     AF071002.1  GI:4704422
KEYWORDS    .
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 732)
AUTHORS     Abbott,G.W., Sesti,F., Splawski,I., Buck,M.E., Lehmann,M.H.,
            Timothy,K.W., Keating,M.T. and Goldstein,S.A.
TITLE       MiRP1 forms IKr potassium channels with HERG and is associated with
            cardiac arrhythmia
JOURNAL     Cell 97 (2), 175-187 (1999)
MEDLINE     99235979
PUBMED      10219939
REFERENCE   2  (bases 1 to 732)
AUTHORS     Abbott,G.W., Sesti,F., Buck,M.E. and Goldstein,S.A.N.
TITLE       Direct Submission
JOURNAL     Submitted (05-JUN-1998) Section of Developmental Biology and
            Biophysics, Department of Pediatrics and Boyer Center for Molecular
            Medicine, Yale University School of Medicine, 295 Congress Avenue,
            New Haven, CT 06536, USA
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q11.1"
                     /tissue_type="heart"
     CDS             74..445
                     /note="potassium channel subunit; MiRP1"
                     /codon_start=1
                     /product="minK-related peptide 1"
                     /protein_id="AAD28086.1"
                     /db_xref="GI:4704423"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"

BASE COUNT      221 a    152 c    157 g    202 t
ORIGIN

Query Match      100.0%;   Score 21; DB 9;   Length 732;
Best Local Similarity  100.0%;  Pred. No. 67;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 ACTTTATCCAATTCACACAG 21
          ||||||||||||||||||||
Db     80 ACTTTATCCAATTCACACAG 100

RESULT 6
AF302095
LOCUS       AF302095                 809 bp    mRNA    linear   PRI 14-SEP-2000
DEFINITION  Homo sapiens voltage-gated K+ channel subunit MIRP1 (KCNE2) mRNA,
            complete cds.
ACCESSION   AF302095
VERSION     AF302095.1  GI:10121887
KEYWORDS    .
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
AUTHORS     Domenech,A., Estivill,X. and de la Luna,S.
TITLE       Cloning of human MIRP1 cDNA
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 809)
AUTHORS     Domenech,A., Estivill,X. and de la Luna,S.
TITLE       Direct Submission
JOURNAL     Submitted (01-SEP-2000) Medical and Molecular Genetics Center,
            Institut Recerca Oncologica, Avia. de Castelldefels Km 2,7,
            L'Hospitalet de Llobregat, Barcelona 08907, Spain
FEATURES             Location/Qualifiers
     source          1..809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22"
     gene            1..809
                     /gene="KCNE2"
     CDS             141..512
                     /gene="KCNE2"
                     /codon_start=1
                     /product="voltage-gated K+ channel subunit MIRP1"
                     /protein_id="AAG13416.1"
                     /db_xref="GI:10121888"
                     /translation="MSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAE
                     NFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                     LEESKATIHENIGAAGFKMSP"

BASE COUNT      247 a    172 c    189 g    200 t    1 others
ORIGIN

Query Match      100.0%;   Score 21; DB 9;   Length 809;
Best Local Similarity  100.0%;  Pred. No. 65;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 ACTTTATCCAATTCACACAG 21
          ||||||||||||||||||||
Db    147 ACTTTATCCAATTCACACAG 167

RESULT 7
AP000320
LOCUS       AP000320               24608 bp    DNA     linear   PRI 20-NOV-1999
DEFINITION  Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
            clone:Q12C8, complete sequence.
ACCESSION   AP000320
VERSION     AP000320.1  GI:4835689
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA, clone:Q12C8.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 24608)
AUTHORS     Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE       Homo sapiens 24,608bp genomic DNA of 21q22.1
JOURNAL     Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 24608)
AUTHORS     Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
            Kitasato Univ. 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     The sequence is a part of the data (ACCESSION No. AP000165 -
            AP000173).
            The sequencing project is supported by Japan Science Technology
```

```
          Corporation (JST) and The Institute of Physical and Chemical
          Research (RIKEN).
FEATURES             Location/Qualifiers
     source          1..24608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="Q12C8"
BASE COUNT       6797 a     5420 c     5437 g     6954 t
ORIGIN

Query Match             100.0%;  Score 21;  DB 9;  Length 24608;
Best Local Similarity   100.0%;  Pred. No. 27;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACTTTATCCAATTCACACAG 21
           ||||||||||||||||||||
Db   15702 ACTTTATCCAATTCACACAG 15722

RESULT 8
AP000052
LOCUS       AP000052      100000 bp      DNA      linear    PRI 20-NOV-1999
DEFINITION  Homo sapiens genomic DNA, chromosome 21q22.1, segment 23/28,
            complete sequence.
ACCESSION   AP000052
VERSION     AP000052.1  GI:3112362
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1998)
REFERENCE   2  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT     This sequence is conducted by Kitasato University JST sequencing
            Laboratory as a JST sequencing team.
            Principal Investigator:Yoshiyuki Sakaki Ph.D.
            Phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
            sakaki@hgc.ims.u-tokyo.ac.jp
            Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D  The
            sequence is submitted by:Human Genome Sequencing in ALIS project of
            JST
            Japan Science and Technology Corporation (JST)
            5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www-alis.tokyo.jst.go.jp/HGS/top.html)
            or send email to webmaster@www-alis.tokyo.jst.go.jp.
FEATURES             Location/Qualifiers
     source          1..100000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="245P17-f4A4f_2"
BASE COUNT      27603 a    21934 c    22513 g    27950 t
ORIGIN

Query Match             100.0%;  Score 21;  DB 9;  Length 100000;
Best Local Similarity   100.0%;  Pred. No. 19;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACTTTATCCAATTCACACAG 21
           ||||||||||||||||||||
Db   80224 ACTTTATCCAATTCACACAG 80244

RESULT 9
AP000167
LOCUS       AP000167      100000 bp      DNA      linear    PRI 08-JAN-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
            clone B2344F14-f50E8, segment 3/9, complete sequence.
ACCESSION   AP000167
VERSION     AP000167.1  GI:4827132
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Homo sapiens 890,291bp genomic DNA of 21q22.1 (REGION: D21S226-AML
            CLONE RANGE: B2344F14-f50E8)
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     E. coli transposon insertion:The present data does not contain E.
            coli transposon sequences which integrated in the
            original/previous sequences.  We determined the boundary between
            the insertion and genomic sequences experimentally, removed the
            insertion sequences, reconstituted the present data. The sequencing
            project is supported by Japan Science Technology Corporation (JST)
            and The Institute of Physical and Chemical Research (RIKEN).
FEATURES             Location/Qualifiers
     source          1..100000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
BASE COUNT      27338 a    22306 c    22571 g    27785 t
ORIGIN

Query Match             100.0%;  Score 21;  DB 9;  Length 100000;
Best Local Similarity   100.0%;  Pred. No. 19;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACTTTATCCAATTCACACAG 21
           ||||||||||||||||||||
Db   72192 ACTTTATCCAATTCACACAG 72212

RESULT 10
AP000120
ID   AP000120    standard; DNA; HUM; 100000 BP.
XX
AC   AP000120;
XX
SV   AP000120.1
XX
DT   04-MAY-1999 (Rel. 59, Created)
DT   26-SEP-1999 (Rel. 61, Last updated, Version 3)
XX
DE   Homo sapiens genomic DNA of 21q22.1, GART and AML related, SLC5A3-f4A4
DE   region, segment 3/8.
XX
KW   HTG.
XX

Qy       1 ACTTTATCCAATTCACACAG 21
```

```
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-100000
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RT   ;
RL   Submitted (15-APR-1999) to the EMBL/GenBank/DDBJ databases.
RL   Mika Hirakawa, Japan Science and Technology Corporation (JST), Advanced
RL   Databases Department; 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081, Japan
RL   (E-mail:mika@tokyo.jst.go.jp, URL:http://www-alis.tokyo.jst.go.jp/,
RL   Tel:81-3-5214-8491, Fax:81-3-5214-8470)
XX
RN   [2]
RA   Hattori M., Ishii K., Toyoda A., Shiba T., Sakaki Y.;
RT   "Homo sapiens 817,199bp genomic DNA of 21q22.1 GART and AML region";
RL   Unpublished.
XX
CC   This sequence is conducted by Kitasato University JST sequencing
CC   Laboratory as a JST sequencing team.
CC   Principal Investigator:Yoshiyuki Sakaki Ph.D.
CC   Phone:+81-3-5449-5622, Fax: +81-3-5449-5445,
CC   sakaki@ghc.u-tokyo.ac.jp
CC   Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D.
CC   The sequence is submitted by Human Genome Sequencing in ALIS
CC   project of JST.
CC   Japan Science and Technology Corporation (JST)
CC   5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081 Japan
CC   For further information about this sequence, including its
CC   location and relationship to other sequences, please visit our
CC   sequence archive Web site (http://www-alis.tokyo.jst.go.jp/HGS/)
CC   or send email to webmaster@www-alis.tokyo.jst.go.jp
XX
FH   Key          Location/Qualifiers
FH
FT   source       1..100000
FT                /chromosome="21"
FT                /db_xref="taxon:9606"
FT                /sequenced_mol="DNA"
FT                /organism="Homo sapiens"
FT                /map="21q22.1"
FT   STS          32213..32320
FT                /note="83F9aL-F/83F9aL-R;The location is between each
FT                flanking site of PCR primers."
FT                /standard_name="D21S1689"
FT   STS          43934..44157
FT                /note="AFMa086yf9;Genethon Marker;The location is between
FT                each flanking site of PCR primers."
FT                /standard_name="D21S1920"
XX
SQ   Sequence 100000 BP; 27731 A; 21990 C; 22220 G; 28059 T; 0 other;

Query Match             100.0%;  Score 21;  DB 17;  Length 100000;
Best Local Similarity   100.0%;  Pred. No. 19;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ACTTTATCCAATTCCACACAG 21
          |||||||||||||||||||||
Db  89526 ACTTTATCCAATTCCACACAG 89546

RESULT 11
AP001719
LOCUS       AP001719      340000 bp    DNA        linear   PRI 10-MAY-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 21q, section 63/105.
ACCESSION   AP001719 AL163264 BA000005
VERSION     AP001719.1  GI:7768719
KEYWORDS    .
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
  TITLE     The DNA sequence of human chromosome 21
  JOURNAL   Nature 405 (6784), 311-319 (2000)
  MEDLINE   20289799
REFERENCE   2  (bases 1 to 340000)
  AUTHORS   Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing
            Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
            Group * Institute of Molecular Biotechnology, Genome Analysis *
            Keio University School of Medicine, Dept. of Molecular Biology *
            GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
            Genetics (addresses see below)
COMMENT     On May 30, 2000 this sequence version replaced gi:7717333.
            The chromosome 21 mapping and sequencing consortium consisting of
            * RIKEN Genomic Sciences Center, Human Genome Research Group, *
            Saganihara 228-8555, Japan,
            * e.mail: hattori@gsc.riken.go.jp
            * URL: http://hgp.gsc.riken.go.jp/
            and
            * Institute of Molecular Biotechnology, Genome Analysis, *
            Beutenbergstrasse 11, D-07745 Jena, Germany,
            * e.mail: gscj-submit@genome.imb-jena.de
            * URL: http://genome.imb-jena.de/
            and
            * Keio University School of Medicine, Molecular Biology, * Tokyo
            160-8582, Japan,
            * e.mail: nshimizu@dmb-med.keio.ac.jp
            * URL: http://www.dmb.med.keio.ac.jp/
            and
            * GBF, Dept. of Genome Analysis,
            * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
            info.genome@gbf.de
            * URL: http://genome.gbf.de/
            and
            * Max-Planck Institute for Molecular Genetics,
            * Ihnestrasse 73, D-14195 Berlin, Germany,
            * e.mail: info-chr21@molgen.mpg.de
            * URL: http://chr21.rz-berlin.mpg.de/
            and
            AL163264: Submitted (10-Apr-2000).
FEATURES             Location/Qualifiers
     source          1..340000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     source          <1..29420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
source              /chromosome="21"
                    /map="21q22.1"
                    /clone="R338L7, 5' partial"
                    /clone_lib="RPCI-11 BAC library"
                    /note="Accession No. AP000467"
                    29242..97350
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="CTD2344F14"
                    /clone_lib="CTD BAC library"
                    /note="Accession No. AP000315"
                    96883..121400
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pS185"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000316"
                    121210..218631
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="P245P17"
                    /clone_lib="RP1 PAC library"
                    /note="Accession No. AP000317"
                    218558..260079
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pQ6F8"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000318"
                    258735..285904
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pQ12C8"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000320"
                    310050..326194
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pQ94A10"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000319"
                    285732..310339
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pQ82F5"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000321"
                    325966..>340000
source              /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="21"
                    /map="21q22.1"
                    /clone="pQ97G8, 3' partial"
                    /clone_lib="CMP21 P1 library"
                    /note="Accession No. AP000322"
repeat_region       complement(1382..1688)
                    /note="AluSp"
                    /rpt_family="SINE/Alu"
                    /rpt_type=DISPERSED
```

```
repeat_region       complement(2099..2390)
                    /note="AluSx"
                    /rpt_family="SINE/Alu"
                    /rpt_type=DISPERSED
repeat_region       2799..2884
                    /note="MIR"
                    /rpt_family="SINE/MIR"
                    /rpt_type=DISPERSED
repeat_region       complement(2891..3287)
                    /note="MER31A"
                    /rpt_family="LTR/MER4-group"
                    /rpt_type=DISPERSED
repeat_region       3288..3381
                    /note="MIR"
                    /rpt_family="SINE/MIR"
                    /rpt_type=DISPERSED
misc_feature        4211..5231
                    /note="CpG island"
misc_feature        4272..4279
                    /note="NotI site"
repeat_region       4328..4476
                    /note="(CGGGG)n"
                    /rpt_family="Simple_repeat"
                    /rpt_type=TANDEM
gene                4392..35127
                    /gene="SLC5A3"
mRNA                join(4392..4740,12254..35127)
                    /gene="SLC5A3"
                    /product="Sodium/myo-inositol cotransporter"
                    /note="human solute carrier family 5, member 3, Accession
                    No. AF027153"
exon                4392..4740
                    /gene="SLC5A3"
                    /number=1
misc_feature        4392..4399
                    /gene="SLC5A3"
                    /note="NotI site"
repeat_region       4810..4985
                    /note="(CGGG)n"
                    /rpt_family="Simple_repeat"
                    /rpt_type=TANDEM
repeat_region       5099..5191
                    /note="(CCG)n"
                    /rpt_family="Simple_repeat"
                    /rpt_type=TANDEM


Query Match             100.0%;    Score 21;  DB 9;  Length 340000;
Best Local Similarity   100.0%;              Pred. No. 14;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACTTTATCCAATTTCACACAG 21
              |||||||||||||||||||||
Db     301433 ACTTTATCCAATTTCACACAG 301453


RESULT 12
AC026481/c
LOCUS       AC026481              99887 bp    DNA        linear   HTG 30-MAR-2000
DEFINITION  Homo sapiens clone RP11-21E3, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC026481
VERSION     AC026481.2  GI:7341959
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99887)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-21E3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 99887)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
```

Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE   Direct Submission
JOURNAL   Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Mar 30, 2000 this sequence version replaced gi:7280305.
           All repeats were identified using RepeatMasker:
           Smit,A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           ------------ Genome Center
                   Center: Whitehead Institute/ MIT Center for Genome Research
                   Center code: WIBR
                   Web site: http://www-seq.wi.mit.edu
                   Contact: sequence submissions@genome.wi.mit.edu
           --------------- Project Information
                   Center project name: L3992
                   Center clone name: 21_E_3
           ----
* NOTE: This record contains 122 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*   1   776: contig of 776 bp in length
*   777   876: gap of   100 bp
*   877 1628: contig of 752 bp in length
*   1629 1728: gap of   100 bp
*   1729 2488: contig of 760 bp in length
*   2489 2588: gap of   100 bp
*   2589 3356: contig of 768 bp in length
*   3357 3456: gap of   100 bp
*   3457 4204: contig of 748 bp in length
*   4205 4304: gap of   100 bp
*   4305 5061: contig of 757 bp in length
*   5062 5161: gap of   100 bp
*   5162 5927: contig of 766 bp in length
*   5928 6027: gap of   100 bp
*   6028 6788: contig of 761 bp in length
*   6789 6888: gap of   100 bp
*   6889 7666: contig of 778 bp in length
*   7667 7766: gap of   100 bp
*   7767 8552: contig of 786 bp in length
*   8553 8652: gap of   100 bp
*   8653 9434: contig of 782 bp in length
*   9435 9534: gap of   100 bp
*   9535 10302: contig of 768 bp in length
*   10303 10402: gap of   100 bp
*   10403 11153: contig of 751 bp in length
*   11154 11253: gap of   100 bp
*   11254 12005: contig of 752 bp in length
*   12006 12105: gap of   100 bp
*   12106 12879: contig of 774 bp in length
*   12880 12979: gap of   100 bp
*   12980 13755: contig of 776 bp in length
*   13756 13855: gap of   100 bp
*   13856 14624: contig of 769 bp in length
*   14625 14724: gap of   100 bp
*   14725 15417: contig of 693 bp in length
*   15418 15517: gap of   100 bp
*   15518 16265: contig of 748 bp in length
*   16266 16365: gap of   100 bp
*   16366 17122: contig of 757 bp in length
*   17123 17222: gap of   100 bp
*   17223 17986: contig of 764 bp in length
*   17987 18086: gap of   100 bp
*   18087 18850: contig of 764 bp in length
*   18851 18950: gap of   100 bp
*   18951 19713: contig of 763 bp in length
*   19714 19813: gap of   100 bp
*   19814 20587: contig of 774 bp in length
*   20588 20687: gap of   100 bp
*   20688 21464: contig of 777 bp in length
*   21465 21564: gap of   100 bp
*   21565 22331: contig of 767 bp in length
*   22332 22431: gap of   100 bp
*   22432 23200: contig of 769 bp in length
*   23201 23300: gap of   100 bp
*   23301 24066: contig of 766 bp in length
*   24067 24166: gap of   100 bp
*   24167 24927: contig of 761 bp in length
*   24928 25027: gap of   100 bp
*   25028 25795: contig of 768 bp in length
*   25796 25895: gap of   100 bp
*   25896 26668: contig of 773 bp in length
*   26669 26768: gap of   100 bp
*   26769 27532: contig of 764 bp in length
*   27533 27632: gap of   100 bp
*   27633 28405: contig of 773 bp in length
*   28406 28505: gap of   100 bp
*   28506 29248: contig of 743 bp in length
*   29249 29348: gap of   100 bp
*   29349 30114: contig of 766 bp in length
*   30115 30214: gap of   100 bp
*   30215 30983: contig of 769 bp in length
*   30984 31083: gap of   100 bp
*   31084 31840: contig of 757 bp in length
*   31841 31940: gap of   100 bp
*   31941 32700: contig of 760 bp in length
*   32701 32800: gap of   100 bp
*   32801 33534: contig of 734 bp in length
*   33535 33634: gap of   100 bp
*   33635 34394: contig of 760 bp in length
*   34395 34494: gap of   100 bp
*   34495 35250: contig of 756 bp in length
*   35251 35350: gap of   100 bp
*   35351 36118: contig of 768 bp in length
*   36119 36218: gap of   100 bp
*   36219 36961: contig of 743 bp in length
*   36962 37061: gap of   100 bp
*   37062 37822: contig of 761 bp in length
*   37823 37922: gap of   100 bp
*   37923 38677: contig of 755 bp in length
*   38678 38777: gap of   100 bp
*   38778 39541: contig of 764 bp in length
*   39542 39641: gap of   100 bp
*   39642 40383: contig of 742 bp in length
*   40384 40483: gap of   100 bp
*   40484 41257: contig of 774 bp in length
*   41258 41357: gap of   100 bp
*   41358 42118: contig of 761 bp in length
*   42119 42218: gap of   100 bp
*   42219 42917: contig of 699 bp in length
*   42918 43017: gap of   100 bp

```
                    /db_xref="GI:21439823"
                    /translation="MSTLSNFAQTLEDVFRRIFITYMDNWRONTTAEQEALQAKVDAE
                    NFYYVILYLMVMIGMFSFIIVALVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILN
                    LEESKATTHENIGAAGFKMSP"
     variation
                    /note="The drug associated here is was Bactrim."
                    95
                    /replace="a"

BASE COUNT      220 a      152 c      158 g      202 t
ORIGIN

Query Match            92.4%;     Score 19.4;  DB 6;  Length 732;
Best Local Similarity  95.2%;     Pred No. 3.3e+02;
Matches    20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1 ACTTTATCCAATTCACACAG 21
        ||||||||||||||||||||| |||||
Db   80 ACTTTATCCAATTCGCACAG 100


RESULT 14
AC078804
LOCUS       AC078804              164456 bp    DNA        linear   HTG 19-JUN-2002
DEFINITION  Homo sapiens chromosome 3 clone RP11-489M17, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC078804
VERSION     AC078804.7  GI:21431061
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164456)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,C., Brown,N.P., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,L.A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,I., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,J., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,J., Newtson,N., Nguyen,N., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Paycon,B., Peery,D., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Soderrgren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
            Usman,I.K., Vasquez,L., Vera,V., Villalon,D., Vinson,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,S., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
```

```
     *   43018   43786: contig of 769 bp in length
     *   43787   43886: gap of    100 bp
     *   43887   44672: contig of 786 bp in length
     *   44673   44772: gap of    100 bp
     *   44773   45538: contig of 766 bp in length
     *   45539   45638: gap of    100 bp
     *   45639   46385: contig of 747 bp in length
     *   46386   46485: gap of    100 bp
     *   46486   47253: contig of 768 bp in length
     *   47254   47353: gap of    100 bp
     *   47354   48093: contig of 740 bp in length
     *   48094   48193: gap of    100 bp
     *   48194   48879: contig of 686 bp in length
     *   48880   48979: gap of    100 bp
     *   48980   49670: contig of 691 bp in length
     *   49671   49770: gap of    100 bp
     *   49771   50451: contig of 681 bp in length
     *   50452   50551: gap of    100 bp
     *   50552   51225: contig of 674 bp in length
     *   51226   51325: gap of    100 bp
     *   51326   52004: contig of 679 bp in length
     *   52005   52104: gap of    100 bp
     *   52105   52767: contig of 663 bp in length
     *   52768   52867: gap of    100 bp
     *   52868   53551: contig of 684 bp in length
     *   53552   53651: gap of    100 bp
     *   53652   54348: contig of 697 bp in length
     *   54349   54448: gap of    100 bp
     *   54449   55136: contig of 688 bp in length
     *   55137   55236: gap of    100 bp
     *   55237   55928: contig of 692 bp in length
     *   55929   56028: gap of    100 bp
     *   56029   56723: contig of 695 bp in length
     *   56724   56823: gap of    100 bp
     *   56824   57520: contig of 697 bp in length
     *   57521   57620: gap of    100 bp
     *   57621   58309: contig of 689 bp in length
     *   58310   58409: gap of    100 bp

Query Match            95.2%;   Score 20; DB 2; Length 99887;
Best Local Similarity  95.2%;   Pred. No. 52;
Matches    20;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    0;

Qy    1 ACTTTATCCAATTTCACACAG 21
        |||||||||||||||||||||
Db 78837 ACTNTATCCAATTTCACACAG 78817


RESULT 13
AX406947
LOCUS       AX406947              732 bp    DNA        linear   PAT 14-JUN-2002
            Sequence 9 from Patent WO0222875.
ACCESSION   AX406947
VERSION     AX406947.1  GI:21439822
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Goldstein,S.A.
  TITLE     Polymorphisms associated with cardiac arrythmia
  JOURNAL   Patent: WO 0222875-A 9 21-MAR-2002;
            YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             74..445
                     /note="minK-related peptide 1, Ala substitued for Thr at
                     amino acid 8"
                     /codon_start=1
                     /protein_id="CAD35512.1"
```

```
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 164456)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (04-AUG-2000) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
REFERENCE    3  (bases 1 to 164456)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (19-JUN-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
COMMENT      On Jun 17, 2002 this sequence version replaced gi:20335729.
             ----------------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Drafting Center Code: BCM
                Contact: hgsc-help@bcm.tmc.edu
             ----------------------- Project Information
                Center project name: HBTN
                Center clone name: RP11-489M17
             ----------------------- Summary Statistics
                Sequencing vector: M13;
                Chemistry: Dye-primer Bodipy: 75% of reads
                Chemistry: Dye-terminator Big Dye: 25% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 158577 bases at least Q40
                Consensus quality: 160039 bases at least Q30
                Consensus quality: 161950 bases at least Q20
                Estimated insert size: 165132; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
             * NOTE: Estimated insert size may differ from sequence length
             *       (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
             * NOTE: This is a 'working draft' sequence. It currently
             *       consists of 7 contigs. The true order of the pieces
             *       is not known and their order in this sequence record is
             *       arbitrary. Gaps between the contigs are represented as
             *       runs of N, but the exact sizes of the gaps are unknown.
             *       This record will be updated with the finished sequence
             *       as soon as it is available and the accession number will
             *       be preserved.
             *          1      7694: contig of 7694 bp in length
             *       7695      7794: gap of unknown length
             *       7795     16159: contig of 8365 bp in length
             *      16160     16259: gap of unknown length
             *      16260     31239: contig of 14980 bp in length
             *      31240     31339: gap of unknown length
             *      31340     52212: contig of 20873 bp in length
             *      52213     52312: gap of unknown length
             *      52313     78975: contig of 26663 bp in length
             *      78976     79075: gap of unknown length
             *      79076    118007: contig of 38932 bp in length
             *     118008    118107: gap of unknown length
             *     118108    164456: contig of 46349 bp in length.
FEATURES             Location/Qualifiers
     source          1..164456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-489M17"
BASE COUNT    52500 a   30465 c   29464 g   50085 t    1942 others
ORIGIN
Query Match     92.4%;  Score 19.4;  DB 2;  Length 164456;
Best Local Similarity   95.2%;  Pred. No. 82;
Matches  20;      Conservative   0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 ACTTTATCCAATTCCACACAG 21
               |||||||||||||||||||||
Db      117867 ACTTTATCCAATTCCACACTG 117887
```

```
 *    3331    4051: contig of 721 bp in length        *   33512   33611: gap of     100 bp
 *    4052    4151: gap of     100 bp                  *   33612   34355: contig of 744 bp in length
 *    4152    4859: contig of 708 bp in length         *   34356   34455: gap of     100 bp
 *    4860    4959: gap of     100 bp                  *   34456   35167: contig of 712 bp in length
 *    4960    5683: contig of 724 bp in length         *   35168   35267: gap of     100 bp
 *    5684    5783: gap of     100 bp                  *   35268   35973: contig of 706 bp in length
 *    5784    6501: contig of 718 bp in length         *   35974   36073: gap of     100 bp
 *    6502    6601: gap of     100 bp                  *   36074   36795: contig of 722 bp in length
 *    6602    7334: contig of 733 bp in length         *   36796   36895: gap of     100 bp
 *    7335    7434: gap of     100 bp                  *   36896   37594: contig of 699 bp in length
 *    7435    8144: contig of 710 bp in length         *   37595   37694: gap of     100 bp
 *    8145    8244: gap of     100 bp                  *   37695   38402: contig of 708 bp in length
 *    8245    8950: contig of 706 bp in length         *   38403   38502: gap of     100 bp
 *    8951    9050: gap of     100 bp                  *   38503   39232: contig of 730 bp in length
 *    9051    9756: contig of 706 bp in length         *   39233   39332: gap of     100 bp
 *    9757    9856: gap of     100 bp                  *   39333   40054: contig of 722 bp in length
 *    9857   10567: contig of 711 bp in length         *   40055   40154: gap of     100 bp
 *   10568   10667: gap of     100 bp                  *   40155   40904: contig of 750 bp in length
 *   10668   11384: contig of 717 bp in length         *   40905   41004: gap of     100 bp
 *   11385   11484: gap of     100 bp                  *   41005   41741: contig of 737 bp in length
 *   11485   12175: contig of 691 bp in length         *   41742   41841: gap of     100 bp
 *   12176   12275: gap of     100 bp                  *   41842   42564: contig of 723 bp in length
 *   12276   12981: contig of 706 bp in length         *   42565   42664: gap of     100 bp
 *   12982   13081: gap of     100 bp                  *   42665   43383: contig of 719 bp in length
 *   13082   13796: contig of 715 bp in length         *   43384   43483: gap of     100 bp
 *   13797   13896: gap of     100 bp                  *   43484   44197: contig of 714 bp in length
 *   13897   14617: contig of 721 bp in length         *   44198   44297: gap of     100 bp
 *   14618   14717: gap of     100 bp                  *   44298   44997: contig of 700 bp in length
 *   14718   15437: contig of 720 bp in length         *   44998   45097: gap of     100 bp
 *   15438   15537: gap of     100 bp                  *   45098   45798: contig of 701 bp in length
 *   15538   16247: contig of 710 bp in length         *   45799   45898: gap of     100 bp
 *   16248   16347: gap of     100 bp                  *   45899   46636: contig of 738 bp in length
 *   16348   17082: contig of 735 bp in length         *   46637   46736: gap of     100 bp
 *   17083   17182: gap of     100 bp                  *   46737   47423: contig of 687 bp in length
 *   17183   17917: contig of 735 bp in length         *   47424   47523: gap of     100 bp
 *   17918   18017: gap of     100 bp                  *   47524   48264: contig of 741 bp in length
 *   18018   18742: contig of 725 bp in length         *   48265   48364: gap of     100 bp
 *   18743   18842: gap of     100 bp                  *   48365   49092: contig of 728 bp in length
 *   18843   19541: contig of 699 bp in length         *   49093   49192: gap of     100 bp
 *   19542   19641: gap of     100 bp                  *   49193   49921: contig of 729 bp in length
 *   19642   20355: contig of 714 bp in length         *   49922   50021: gap of     100 bp
 *   20356   20455: gap of     100 bp                  *   50022   50749: contig of 728 bp in length
 *   20456   21186: contig of 731 bp in length         *   50750   50849: gap of     100 bp
 *   21187   21286: gap of     100 bp                  *   50850   51561: contig of 712 bp in length
 *   21287   21983: contig of 697 bp in length         *   51562   51661: gap of     100 bp
 *   21984   22083: gap of     100 bp                  *   51662   52375: contig of 714 bp in length
 *   22084   22806: contig of 723 bp in length         *   52376   52475: gap of     100 bp
 *   22807   22906: gap of     100 bp                  *   52476   53198: contig of 723 bp in length
 *   22907   23625: contig of 719 bp in length         *   53199   53298: gap of     100 bp
 *   23626   23725: gap of     100 bp                  *   53299   54016: contig of 718 bp in length
 *   23726   24463: contig of 738 bp in length         *   54017   54116: gap of     100 bp
 *   24464   24563: gap of     100 bp                  *   54117   54824: contig of 708 bp in length
 *   24564   25297: contig of 734 bp in length         *   54825   54924: gap of     100 bp
 *   25298   25397: gap of     100 bp                  *   54925   55651: contig of 727 bp in length
 *   25398   26110: contig of 713 bp in length         *   55652   55751: gap of     100 bp
 *   26111   26210: gap of     100 bp                  *   55752   56460: contig of 709 bp in length
 *   26211   26924: contig of 714 bp in length         *   56461   56560: gap of     100 bp
 *   26925   27024: gap of     100 bp                  *   56561   57299: contig of 739 bp in length
 *   27025   27731: contig of 707 bp in length
 *   27732   27831: gap of     100 bp
 *   27832   28552: contig of 721 bp in length
 *   28553   28652: gap of     100 bp
 *   28653   29364: contig of 712 bp in length
 *   29365   29464: gap of     100 bp
 *   29465   30201: contig of 737 bp in length
 *   30202   30301: gap of     100 bp
 *   30302   31050: contig of 749 bp in length
 *   31051   31150: gap of     100 bp
 *   31151   31847: contig of 697 bp in length
 *   31848   31947: gap of     100 bp
 *   31948   32684: contig of 737 bp in length
 *   32685   32784: gap of     100 bp
 *   32785   33511: contig of 727 bp in length
```

```
Query Match            87.6%;  Score 18.4;  DB 2;  Length 64677;
Best Local Similarity  95.0%;  Pred. No. 2.8e+02;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy       1 ACTTATCCAATTCACACA 20
           || |||||||||||||||
Db   49052 ACATTATCCAATTCACACA 49033
```

Search completed: June 9, 2003, 09:09:53
Job time : 619 secs

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 08:36:49 ; Search time 119 Seconds
                                    (without alignments)
                           397.411 Million cell updates/sec

Title:           US-09-550-163-1_COPY_80_100
Perfect score:   21
Sequence:        1 acttatccaattcacacag 21

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:    4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  N_Geneseq_101002:*
       1:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
       2:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
       3:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
       4:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
       5:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
       6:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
       7:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
       8:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
       9:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
      10:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
      11:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
      12:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
      13:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
      14:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
      15:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
      16:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
      17:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
      18:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
      19:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
      20:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
      21:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
      22:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
      23:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
      24:   /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	
1	21	100.0	372	22	AAI24432		Probe #14365 for g
2	21	100.0	372	22	AAI09965		Probe #9956 used t
3	21	100.0	372	22	AAS00245		Human potassium ch
4	21	100.0	450	22	ABA44797		Human breast cell
5	21	100.0	450	22	ABA55252		Human foetal liver
6	21	100.0	450	22	ABA24997		Probe #3463 for ge
7	21	100.0	450	22	AAK03508		Human brain expres
8	21	100.0	450	22	AAK28962		Human bone marrow
9	21	100.0	450	22	AAI13549		Probe #3482 for ge
10	21	100.0	450	22	AAI34911		Probe #3597 used t
11	21	100.0	450	22	AAI03438		Probe #3429 used t
12	21	100.0	450	24	ABS03496		Human genome-deriv
13	21	100.0	471	22	AAI15256		Probe #5189 for ge
14	21	100.0	471	22	AAI04990		Probe #4981 used t
15	21	100.0	471	22	AAF80269		Nucleotide sequenc
16	21	100.0	600	22	ABA09192		Human MiRP1 homolo
17	21	100.0	600	22	AAK52645		Human polynucleoti
18	21	100.0	655	22	AAK51661		Human polynucleoti
19	21	100.0	732	21	AAC64071		Human potassium ch
20	21	100.0	732	21	AAC64084		Human potassium ch
21	21	100.0	732	21	AAC64085		Human potassium ch
22	21	100.0	732	24	ABK86573		cDNA encoding huma
23	21	100.0	732	24	AAD35169		Human KCNE2 mutant
24	21	100.0	732	24	AAD35170		Human KCNE2 wild t
25	21	100.0	732	24	AAD35171		Human KCNE2 mutant
26	21	100.0	732	21	AAD35172		Human KCNE2 mutant
27	19.4	92.4	732	21	AAC64083		Human potassium ch
28	19.4	92.4	732	21	AAC64086		Human potassium ch
29	19.4	92.4	732	24	AAD35173		Human KCNE2 mutant
30	17.8	84.8	877	23	ABV10075		Human prostate exp
31	17.8	84.8	2481	22	AAK74863	c	Human immune/haema
32	17.8	84.8	21332	22	AAK74878	c	Human immune/haema
33	17.8	84.8	21332	22	AAK83177	c	Human immune/haema
34	17.4	82.9	1038602	20	AAZ01425		Complete genome se
35	17	81.0	2384	22	AAH17956		Human cDNA sequenc
36	16.8	80.0	7800	23	ABL24460		Drosophila melanog
37	16.8	80.0	8342	24	ABL32501	c	Human immune syste
38	16.8	80.0	15987	23	ABL14198		Drosophila melanog
39	16.4	78.1	264	24	AAI72099		Anergy related cDN
40	16.4	78.1	297	24	ABQ67893		Listeria monocytog
41	16.4	78.1	297	24	ABQ69916		Listeria monocytog
42	16.4	78.1	322	22	AAK69409		Human immune/haema
43	16.4	78.1	322	22	AAK69411		Human immune/haema
44	16.4	78.1	339	24	ABL86737		Human ovarian canc
45	16.4	78.1	585	23	ABV53743		Human prostate exp

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
AAI24432
ID   AAI24432 standard; DNA; 372 BP.
XX
AC   AAI24432;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #14365 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO20015278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0612366.
PR   21-SEP-2000; 2000US-0214687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
```

XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 14365; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%;   Score 21; DB 22; Length 372;
Best Local Similarity 100.0%;   Pred. No. 6.1;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTCACACAG 21
             |||||||||||||||||||||
Db   7 ACTTTATCCAATTCACACAG 27

RESULT 2
AAI09965
ID   AAI09965 standard; DNA; 372 BP.
XX
AC   AAI09965;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #9956 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 9956; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe

CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match   100.0%;   Score 21; DB 22; Length 372;
Best Local Similarity 100.0%;   Pred. No. 6.1;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTCACACAG 21
             |||||||||||||||||||||
Db   7 ACTTTATCCAATTCACACAG 27

RESULT 3
AAS00245
ID   AAS00245 standard; DNA; 372 BP.
XX
AC   AAS00245;
XX
DT   10-MAY-2001 (first entry)
XX
DE   Human potassium channel regulatory protein, Mink2, DNA sequence.
XX
KW   Human; Mink2; potassium channel; cardiac arrhythmia; hypertension; ds;
KW   angina; asthma; diabetes; renal insufficiency; urinary incontinence;
KW   irritable colon; epilepsy; cerebrovascular ischaemia; autoimmune disease.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..372
FT   /tag= a
FT   /product= "MINK2 potassium channel protein"
XX
PN   WO20011403-A1.
XX
PD   01-MAR-2001.
XX
PF   18-AUG-2000; 2000WO-US22799.
XX
PR   20-AUG-1999; 99US-0379201.
XX
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI   Ficker E, Wible B, Brown AM;
XX
DR   WPI; 2001-218424/22.
DR   P-PSDB; AAU00215.
XX
PT   Novel potassium channel gene termed MinK2 encoding potassium channel
PT   regulatory protein, useful for screening compounds that are useful for
PT   treating diseases caused by aberrant potassium activity -
XX
PS   Claim 1; Fig 9; 39pp; English.
XX
CC   The sequence represents the coding sequence of human potassium channel
CC   regulatory protein, Mink2. Mink2 sequence is useful for producing a
CC   potassium channel regulatory protein useful for in vitro or in vivo
CC   screening of agonistic or antagonistic compounds that are useful for
CC   treating diseases caused by aberrant potassium activity, such as human
CC   cardiac arrhythmias, hypertension, angina, asthma, diabetes, renal
CC   insufficiency, urinary incontinence, irritable colon, epilepsy,
CC   cerebrovascular ischaemia, and autoimmune disease.
XX

```
SQ   Sequence 372 BP; 110 A; 90 C; 82 G; 90 T; 0 other;

Query Match             100.0%;  Score 21;  DB 22;  Length 372;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ACTTTATCCAATTCACACAG 21
            |||||||||||||||||||||
Db        7 ACTTTATCCAATTCACACAG 27

RESULT 4
ABA44797
ID   ABA44797 standard; DNA; 450 BP.
XX
AC   ABA44797;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3492.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes -
XX
PS   Claim 1; SEQ ID NO 3492; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;


Query Match             100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ACTTTATCCAATTCACACAG 21
            |||||||||||||||||||||
Db      162 ACTTTATCCAATTCACACAG 182

RESULT 5
ABA55252
ID   ABA55252 standard; DNA; 450 BP.
XX
AC   ABA55252;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3557.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 3557; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;


Query Match             100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ACTTTATCCAATTCACACAG 21
            |||||||||||||||||||||
Db      162 ACTTTATCCAATTCACACAG 182

RESULT 6
ABA24997
ID   ABA24997 standard; DNA; 450 BP.
XX
```

```
AC   ABA24997;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Probe #3463 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 3463; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match             100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACTTTATCCAATTTCACACAG 21
              |||||||||||||||||||||
Db        162 ACTTTATCCAATTTCACACAG 182

RESULT 7
AAK03508
ID   AAK03508 standard; DNA; 450 BP.
XX
AC   AAK03508;
XX
DT   05-NOV-2001 (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 3499.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
```

```
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0612366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 3499; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match             100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACTTTATCCAATTTCACACAG 21
              |||||||||||||||||||||
Db        162 ACTTTATCCAATTTCACACAG 182

RESULT 8
AAK28962
ID   AAK28962 standard; DNA; 450 BP.
XX
AC   AAK28962;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3519.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0633366.
```

```
PR    21-SEP-2000; 2000US-0234687.
PR    27-SEP-2000; 2000US-0236359.
PR    04-OCT-2000; 2000GB-0024263.
XX
PA    (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI    Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR    WPI; 2001-488900/53.
XX
PT    Human genome-derived single exon nucleic acid probes useful for
PT    analyzing gene expression in human bone marrow -
XX
PS    Example 4; SEQ ID NO: 3519; 658pp + Sequence Listing; English.
XX
CC    The present invention provides a number of single exon nucleic acid
CC    probes which are derived from genomic sequences expressed in the human
CC    bone marrow. They can be used to measure gene expression in bone marrow
CC    samples, which may enable the improved diagnosis and treatment of cancers
CC    such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC    the probes of the invention.
XX
SQ    Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match      100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ACTTTATCCAATTCACACAG 21
          ||||||||||||||||||||
Db   162 ACTTTATCCAATTCACACAG 182

RESULT 9
AAI13549
ID   AAI13549 standard; DNA; 450 BP.
XX
AC   AAI13549;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #3482 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.          .
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0612366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 3482; 487pp; English.
XX
```

```
CC    The present invention relates to human single exon nucleic acid probes
CC    (SENP). The present sequence is one such probe. The SENPs are derived
CC    from human HeLa cells. The SENPs can be used to produce a single exon
CC    microarray, which can be used for measuring human gene expression in a
CC    sample derived from human cervical epithelial cells. By measuring gene
CC    expression, the probes are therefore useful in grading and/or staging
CC    of diseases of the cervix, notably cervical cancer.
CC    Note: The sequence data for this patent did not form part of the printed
CC    specification, but was obtained in electronic format directly from WIPO
CC    at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ    Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match      100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ACTTTATCCAATTCACACAG 21
          ||||||||||||||||||||
Db   162 ACTTTATCCAATTCACACAG 182

RESULT 10
AAI34911
ID   AAI34911 standard; DNA; 450 BP.
XX
AC   AAI34911;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #3597 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 3597; 654pp; English.
XX
CC    The present invention relates to single exon nucleic acid probes (SENP).
CC    The present sequence is one such probe. The probes are useful for
CC    producing a microarray for predicting, measuring and displaying gene
CC    expression in samples derived from human placenta. The probes are useful
CC    for antenatal diagnosis of human genetic disorders.
XX
SQ    Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match      100.0%;  Score 21;  DB 22;  Length 450;
Best Local Similarity   100.0%;  Pred. No. 6.1;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy          1 ACTTTATCCAATTTCACACAG 21
              ||||||||||||||||||||
Db        162 ACTTTATCCAATTTCACACAG 182

RESULT 11
AAI03438
ID   AAI03438 standard; DNA; 450 BP.
XX
AC   AAI03438;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #3429 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 3429; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match      100.0%;   Score 21;   DB 22;   Length 450;
Best Local Similarity  100.0%;   Pred. No. 6.1;
Matches   21;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy          1 ACTTTATCCAATTTCACACAG 21
              ||||||||||||||||||||
Db        162 ACTTTATCCAATTTCACACAG 182

RESULT 12
ABS03496
ID   ABS03496 standard; DNA; 450 BP.
XX
AC   ABS03496;
XX
DT   19-AUG-2002 (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 3487.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 1; SEQ ID No 3487; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
```

CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe of the invention.
CC   Note:The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 450 BP; 136 A; 110 C; 86 G; 118 T; 0 other;

Query Match   100.0%; Score 21; DB 24; Length 450;
Best Local Similarity   100.0%; Pred. No. 6.1;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAG 21
              |||||||||||||||||||||
Db   162 ACTTTATCCAATTTCACACAG 182

RESULT 13
AAI15256
ID   AAI15256 standard; DNA; 471 BP.
XX
AC   AAI15256;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #5189 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 5189; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO

CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 471 BP; 148 A; 102 C; 86 G; 135 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 471;
Best Local Similarity   100.0%; Pred. No. 6.1;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAG 21
              |||||||||||||||||||||
Db   245 ACTTTATCCAATTTCACACAG 265

RESULT 14
AAI04990
ID   AAI04990 standard; DNA; 471 BP.
XX
AC   AAI04990;
XX
DT   09-OCT-2001 (first entry)
XX
DE   Probe #4981 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 4981; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridies at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 471 BP; 148 A; 102 C; 86 G; 135 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 471;
Best Local Similarity   100.0%; Pred. No. 6.1;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAG 21
               |||||||||||||||||||||
Db   245 ACTTTATCCAATTTCACACAG 265

RESULT 15
AAF80269
ID   AAF80269 standard; DNA; 471 BP.
XX
AC   AAF80269;
XX
DT   29-JUN-2001 (first entry)
XX
DE   Nucleotide sequence of human potassium channel subunit IsK2.
XX
KW   Human; potassium channel; IsK2; gene therapy; gastric motility;
KW   gastric acid secretion; anti-arrhythmic agent; myocardial infarction; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   79..450
FT   /*tag= a
FT   /product= "potassium channel subunit IsK2"
XX
PN   WO20127246-A1.
XX
PD   19-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US28014.
XX
PR   12-OCT-1999; 99US-0158781.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Swanson RJ, Liu Y, Folander K;
XX
DR   WPI; 2001-273764/28.
DR   P-PSDB; AAB67800.
XX
PT   New DNA encoding the IsK2 potassium channel subunit, useful e.g. for
PT   detecting mutations and screening for therapeutic agents -
XX
PS   Claim 3; Fig 1A; 46pp; English.
XX
CC   The present sequence encodes a human potassium channel subunit,
CC   designated IsK2. The IsK2 polynucleotide, and derived probes, are
CC   used diagnostically to detect mutations in the IsK2 gene, to determine
CC   levels of mRNA expression and to isolate homologous sequences; for
CC   recombinant expression of IsK2; in gene therapy to increase potassium
CC   channel activity and to generate transgenic animals, as models and
CC   for drug screening. Recombinant IsK2 is used for studying biochemical
CC   activity of IsK2 and its role in disorders of gastric motility and
CC   gastric acid secretion, and to raise specific antibodies. IsK2
CC   modulators are potentially useful for treating diseases associated with
CC   increased or reduced potassium channel activity, e.g. as
CC   anti-arrhythmic agents for treating myocardial infarction and as
CC   regulators of gastric acid secretion.
XX
SQ   Sequence 471 BP; 143 A; 110 C; 103 G; 115 T; 0 other;

Query Match   100.0%; Score 21; DB 22; Length 471;
Best Local Similarity   100.0%; Pred. No. 6.1;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTTTATCCAATTTCACACAG 21
                |||||||||||||||||||||
Db   85 ACTTTATCCAATTTCACACAG 105

Search completed: June 9, 2003, 08:51:24
Job time : 121 secs

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     June  9, 2003, 08:36:51 ; Search time 31.5 Seconds
                                       (without alignments)
                                       204.451 Million cell updates/sec

Title:           US-09-550-163-1_COPY_80_100
Perfect score:   21
Sequence:        1 acttatccaattcacacag 21

Scoring table:   IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:        882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued Patents NA:*
                 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
c 1	16.4	78.1	6402	2	US-08-670-707A-36	Sequence 36, Appl
c 2	16.4	78.1	6402	4	US-09-037-601-36	Sequence 36, Appl
c 3	16.4	78.1	6402	4	US-09-315-179-36	Sequence 36, Appl
c 4	16.4	78.1	6402	4	US-09-523-656-29	Sequence 29, Appl
c 5	15.8	75.2	627	4	US-08-962-281-1	Sequence 1, Appl
c 6	15.8	75.2	1211	4	US-09-199-603C-25	Sequence 25, Appl
c 7	15.4	75.2	5000	2	US-09-147-522-5	Sequence 5, Appli
c 8	15.4	73.3	2244	2	US-08-203-532F-1	Sequence 1, Appli
c 9	15.4	73.3	2244	4	US-09-078-465-1	Sequence 1, Appli
c 10	15.4	73.3	2244	5	PCT-US95-01882A-1	Sequence 1, Appli
c 11	15.4	73.3	2430	4	US-08-845-258-3	Sequence 3, Appli
c 12	15.4	73.3	2430	4	US-08-845-258-40	Sequence 40, Appli
c 13	15.4	73.3	2430	4	US-08-990-571-3	Sequence 3, Appli
c 14	15.4	73.3	2430	4	US-08-990-571-40	Sequence 40, Appli
c 15	15.4	73.3	2430	4	US-08-723-142A-3	Sequence 3, Appli
c 16	15.4	73.3	2430	4	US-08-723-142A-40	Sequence 40, Appli
c 17	15.4	73.3	2430	4	US-09-528-784A-3	Sequence 3, Appli
c 18	15.4	73.3	2430	4	US-09-528-784A-40	Sequence 40, App
c 19	15.4	73.3	43360	4	US-09-453-702B-206	Sequence 206, App
c 20	15.4	73.3	45325	4	US-09-453-702B-261	Sequence 261, App
c 21	15.2	72.4	865	4	US-09-280-116-63	Sequence 63, Appli
c 22	15.2	72.4	865	4	US-09-796-110-1	Sequence 1, Appli
c 23	15.2	72.4	1001	4	US-09-641-638-441	Sequence 441, App
c 24	15.2	72.4	1001	4	US-09-641-638-450	Sequence 450, App
c 25	15.2	72.4	1622	4	US-09-004-393B-1	Sequence 1, Appli
c 26	15.2	72.4	1713	4	US-08-759-436-1	Sequence 1, Appli
c 27	15.2	72.4	2706	2	US-08-630-822A-61	Sequence 61, Appl
c 28	15.2	72.4	2706	2	US-09-005-069-61	Sequence 61, Appl
c 29	15.2	72.4	2706	4	US-09-171-156A-20	Sequence 20, Appl
30	15.2	72.4	4629	4	US-08-484-891-7	Sequence 7, Appli
31	15.2	72.4	4670	3	US-08-717-294-41	Sequence 41, Appl
32	15.2	72.4	4999	4	US-09-470-618-14	Sequence 14, Appl
33	15.2	72.4	4999	4	US-09-364-862-14	Sequence 14, Appl
34	15.2	72.4	5035	2	US-08-882-083-1	Sequence 1, Appli
35	15.2	72.4	5035	3	US-08-558-107-1	Sequence 1, Appli
36	15.2	72.4	5035	3	US-09-243-539-1	Sequence 1, Appli
37	15.2	72.4	6999	1	US-08-276-594A-1	Sequence 1, Appli
38	15.2	72.4	7056	1	US-08-121-202-1	Sequence 1, Appli
39	15.2	72.4	8241	6	5171844-1	Patent No. 5171844
40	15.2	72.4	8967	1	US-08-366-851A-1	Sequence 1, Appli
41	15.2	72.4	9009	1	US-07-864-004B-3	Sequence 3, Appli
42	15.2	72.4	9009	1	US-08-251-937A-3	Sequence 3, Appli
43	15.2	72.4	9009	1	US-08-212-133A-1	Sequence 1, Appli
44	15.2	72.4	9009	1	US-08-474-503-1	Sequence 1, Appli
45	15.2	72.4	9009	2	US-08-670-707A-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-670-707A-36/c
; Sequence 36, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
;   APPLICANT: Lollar, John S.
;   TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;   NUMBER OF SEQUENCES: 40
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;     STREET: 5370 Manhattan Circle Suite 201
;     CITY: Boulder
;     STATE: Colorado
;     COUNTRY: USA
;     ZIP: 80303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/670,707A
;     FILING DATE: 26-JUN-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: WO PCT/US94/13200
;     FILING DATE: 15-NOV-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/212,133
;     FILING DATE: 11-MAR-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/864,004
;     FILING DATE: 07-APR-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Greenlee, Lorance L.
;     REGISTRATION NUMBER: 27,894
;     REFERENCE/DOCKET NUMBER: 75-95F
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 303/499-8080
;     TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6402 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: not relevant
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;     ORGANISM: Pig
```

```
  FEATURE:
;    NAME/KEY: CDS
;    LOCATION: 1..6402
US-08-670-707A-36

Query Match             78.1%;  Score 16.4;  DB 2;  Length 6402;
Best Local Similarity   94.4%;  Pred. No. 40;
  Matches   17;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy          3 TTTATCCAATTCACACA 20
              ||||||||||||| |||||
Db       3116 TTTATCCAATTGACACA 3099


RESULT 2
US-09-037-601-36/c
; Sequence 36, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
;   APPLICANT: Lollar, John S.
;   TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;   NUMBER OF SEQUENCES: 40
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;     STREET: 5370 Manhattan Circle Suite 201
;     CITY: Boulder
;     STATE: Colorado
;     COUNTRY: USA
;     ZIP: 80303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/037,601
;     FILING DATE: 26-JUN-1996
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: WO PCT/US94/13200
;     FILING DATE: 15-NOV-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/212,133
;     FILING DATE: 11-MAR-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/864,004
;     FILING DATE: 07-APR-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Ferber, Donna M.
;     REGISTRATION NUMBER: 33,878
;     REFERENCE/DOCKET NUMBER: 75-95F
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 303/499-8080
;     TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6402 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: not relevant
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;     ORGANISM: Pig
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..6402
US-09-037-601-36

Query Match             78.1%;  Score 16.4;  DB 4;  Length 6402;
Best Local Similarity   94.4%;  Pred. No. 40;
  Matches   17;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;
```

```
Qy          3 TTTATCCAATTCACACA 20
              ||||||||||||| |||||
Db       3116 TTTATCCAATTGACACA 3099


RESULT 3
US-09-315-179-36/c
; Sequence 36, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
;   APPLICANT: Lollar, John S
;   TITLE OF INVENTION: Modified Factor VIII
;   FILE REFERENCE: 75-95H
;   CURRENT APPLICATION NUMBER: US/09/315,179
;   CURRENT FILING DATE: 1999-05-20
;   EARLIER APPLICATION NUMBER: U.S. 09/037,601
;   EARLIER FILING DATE: 1998-03-10
;   EARLIER APPLICATION NUMBER: U.S. 08/670,707
;   EARLIER FILING DATE: 1996-06-26
;   EARLIER APPLICATION NUMBER: PCT/US97/11155
;   EARLIER FILING DATE: 1997-06-26
;   EARLIER APPLICATION NUMBER: PCT/US94/13200
;   EARLIER FILING DATE: 1994-11-15
;   EARLIER APPLICATION NUMBER: U.S. 08/212,133
;   EARLIER FILING DATE: 1994-03-11
;   EARLIER APPLICATION NUMBER: U.S. 07/864,004
;   EARLIER FILING DATE: 1992-04-07
;   NUMBER OF SEQ ID NOS: 40
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 6402
;   TYPE: DNA
;   ORGANISM: Sus scrofa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(6399)
US-09-315-179-36

Query Match             78.1%;  Score 16.4;  DB 4;  Length 6402;
Best Local Similarity   94.4%;  Pred. No. 40;
  Matches   17;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy          3 TTTATCCAATTCACACA 20
              ||||||||||||| |||||
Db       3116 TTTATCCAATTGACACA 3099


RESULT 4
US-09-523-656-29/c
; Sequence 29, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
;   APPLICANT: Lollar S., John
;   TITLE OF INVENTION: MODIFIED FACTOR VIII
;   FILE REFERENCE: 75-95I
;   CURRENT APPLICATION NUMBER: US/09/523,656
;   CURRENT FILING DATE: 2000-03-10
;   EARLIER APPLICATION NUMBER: 09/037,601
;   EARLIER FILING DATE: 1998-03-10
;   EARLIER APPLICATION NUMBER: 08/670,707
;   EARLIER FILING DATE: 1996-06-26
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 6402
;   TYPE: DNA
;   ORGANISM: Porcine
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(6399)
US-09-523-656-29

Query Match             78.1%;  Score 16.4;  DB 4;  Length 6402;
```

Best Local Similarity 94.4%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      3 TTTATCCAATTTCACACA 20
          |||||||| |||||
Db   3116 TTTATCCAATTTGACACA 3099
```

RESULT 5
US-08-962-281-1/c
; Sequence 1, Application US/08962281
; Patent No. 6403356
; GENERAL INFORMATION:
; APPLICANT: You, Li
; APPLICANT: Usher, John J.
; APPLICANT: White, Brenda J.
; APPLICANT: No. 6403356otny, Jiri
; TITLE OF INVENTION: MUTANT PENICILLIN G ACYLSES
; FILE REFERENCE: ON0149A (9197F-86)
; CURRENT APPLICATION NUMBER: US/08/962,281
; CURRENT FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(627)
US-08-962-281-1

Query Match 75.2%; Score 15.8; DB 4; Length 627;
Best Local Similarity 89.5%; Pred. No. 57;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy     2 CTTTATCCAATTTCACACA 20
          ||||||||||||||
Db   202 CTTTATCAAATTTCACAAA 184
```

RESULT 6
US-09-198-603C-25
; Sequence 25, Application US/09198603C
; Patent No. 6337193
; GENERAL INFORMATION:
; APPLICANT: TULLY, Raymond E.
; APPLICANT: CALTAGIRONE, G. Thomas
; APPLICANT: MOYER, Shawn S.
; APPLICANT: RONNING, Michael T.
; TITLE OF INVENTION: EXPRESSION OF MANNOSE-BINDING PROTEIN IN METHYLOTROPHIC
; TITLE OF INVENTION: YEAST
; FILE REFERENCE: A7290
; CURRENT APPLICATION NUMBER: US/09/198,603C
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 1211
;   TYPE: DNA
;   ORGANISM: Mouse
US-09-198-603C-25

Query Match 75.2%; Score 15.8; DB 4; Length 1211;
Best Local Similarity 89.5%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy     3 TTTATCCAATTTCACACAG 21
          ||||||||||||||
Db    56 TTTATTCAATTTCACAGAG 74
```

RESULT 7

US-09-147-522-5
; Sequence 5, Application US/09147522
; Patent No. 6107069
; GENERAL INFORMATION:
; APPLICANT: MAGAGNIN, SIMONA
; APPLICANT: BENATTI, LUCA
; APPLICANT: CINI, MASSIMO
; APPLICANT: SPECIALE, CARMELA
; APPLICANT: COVINI, NEVIE
; TITLE OF INVENTION: RECOMBINANT KYNURENINE-3-HYDROXYLASE ENZYME AND
; TITLE OF INVENTION: PROCESS FOR ITS PREPARATION
; FILE REFERENCE: 0769-0408-OPCT
; CURRENT APPLICATION NUMBER: US/09/147,522
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: PCT/EP7/03589
; EARLIER FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 5000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (47)..(1507)
US-09-147-522-5

Query Match 75.2%; Score 15.8; DB 3; Length 5000;
Best Local Similarity 89.5%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ACTTTATCCAATTTCACAC 19
          |||||| |||||||| ||
Db   2587 ACTTTATTCAATTTCAGAC 2605
```

RESULT 8
US-08-203-532F-1/c
; Sequence 1, Application US/08203532F
; Patent No. 5856121
; GENERAL INFORMATION:
; APPLICANT: Gorski, David H.
; APPLICANT: Walsh, Kenneth
; TITLE OF INVENTION: Growth Arrest Homeobox Gene
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calfee, Halter, and Griswold
;   STREET: 800 Superior Avenue
;   CITY: Cleveland
;   STATE: Ohio
;   COUNTRY: U.S.A.
;   ZIP: 44114-2688
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/203,532F
;   FILING DATE:
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Golrick, Mary E.
;   REGISTRATION NUMBER: 34829
;   REFERENCE/DOCKET NUMBER: 22311/00114
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (216) 622-8200
;   TELEFAX: (216) 241-0816
;   TELEX: 980499
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2244 base pairs
;   TYPE: nucleic acid

```
        STRANDEDNESS: both
        TOPOLOGY: linear
        MOLECULE TYPE: cDNA
        HYPOTHETICAL: NO
        ANTI-SENSE: NO
        FEATURE:
          NAME/KEY: CDS
          LOCATION: 197..1108
US-08-203-532F-1

Query Match     73.3%;  Score 15.4;  DB 2;  Length 2244;
Best Local Similarity  94.1%;  Pred. No. 1e+02;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy      1 ACTTTATCCAATTTCAC 17
          ||||| |||||||||||
Db   1314 ACTTTCTCCAATTTCAC 1298

RESULT 9
US-09-078-465-1/c
; Sequence 1, Application US/09078465
; Patent No. 6280969
; GENERAL INFORMATION:
;   APPLICANT: Gorski, David H.
;   APPLICANT: Walsh, Kenneth
;   TITLE OF INVENTION: Growth Arrest Homeobox Gene
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Calfee, Halter, and Griswold
;     STREET: 800 Superior Avenue
;     CITY: Cleveland
;     STATE: Ohio
;     COUNTRY: U.S.A.
;     ZIP: 44114-2688
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/078,465
;     FILING DATE: 14-MAY-1998
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/08/203,532
;     FILING DATE: 24-FEB-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Golrick, Mary E.
;     REGISTRATION NUMBER: 34829
;     REFERENCE/DOCKET NUMBER: 22311/00114
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (216) 622-8200
;     TELEFAX: (216) 241-0816
;     TELEX: 980499
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2244 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: both
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 197..1108
US-09-078-465-1

Query Match     73.3%;  Score 15.4;  DB 4;  Length 2244;
Best Local Similarity  94.1%;  Pred. No. 1e+02;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;


Qy      1 ACTTTATCCAATTTCAC 17
          ||||| |||||||||||
Db   1314 ACTTTCTCCAATTTCAC 1298

RESULT 10
PCT-US95-01882A-1/c
; Sequence 1, Application PC/TUS9501882A
; GENERAL INFORMATION:
;   APPLICANT: Gorski, David H.
;   APPLICANT: Walsh, Kenneth
;   TITLE OF INVENTION: Growth Arrest Homeobox Gene
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Calfee, Halter, and Griswold
;     STREET: 800 Superior Avenue
;     CITY: Cleveland
;     STATE: Ohio
;     COUNTRY: U.S.A.
;     ZIP: 44114-2688
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US95/01882A
;     FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Golrick, Mary E.
;     REGISTRATION NUMBER: 34829
;     REFERENCE/DOCKET NUMBER: 22311/00114
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (216) 622-8200
;     TELEFAX: (216) 241-0816
;     TELEX: 980499
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2244 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: both
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 197..1108
PCT-US95-01882A-1

Query Match     73.3%;  Score 15.4;  DB 5;  Length 2244;
Best Local Similarity  94.1%;  Pred. No. 1e+02;
Matches   16;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy      1 ACTTTATCCAATTTCAC 17
          ||||| |||||||||||
Db   1314 ACTTTCTCCAATTTCAC 1298

RESULT 11
US-08-845-258-3
; Sequence 3, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Lodes, Michael J.
;   APPLICANT: Houghton, Raymond
;   APPLICANT: Sleath, Paul R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;   TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
;   NUMBER OF SEQUENCES: 53
;   CORRESPONDENCE ADDRSS:
```

```
               ADDRESSEE: SEED AND BERRY
               STREET: 6300 Columbia Center, 701 Fifth Avenue
               CITY: Seattle
               STATE: Washington
               COUNTRY: USA
               ZIP: 98104
           COMPUTER READABLE FORM:
               MEDIUM TYPE: Floppy disk
               COMPUTER: IBM PC compatible
               OPERATING SYSTEM: PC-DOS/MS-DOS
               SOFTWARE: PatentIn Release #1.0, Version #1.30
           CURRENT APPLICATION DATA:
               APPLICATION NUMBER: US/08/845,258
               FILING DATE: 24-APR-1997
               CLASSIFICATION: 435
           ATTORNEY/AGENT INFORMATION:
               NAME: Maki, David J.
               REGISTRATION NUMBER: 31,392
               REFERENCE/DOCKET NUMBER: 210121.426C1
           TELECOMMUNICATION INFORMATION:
               TELEPHONE: (206) 622-4900
               TELEFAX: (206)682-6031
       INFORMATION FOR SEQ ID NO: 3:
           SEQUENCE CHARACTERISTICS:
               LENGTH: 2430 base pairs
               TYPE: nucleic acid
               STRANDEDNESS: single
               TOPOLOGY: linear
US-08-845-258-3

  Query Match      73.3%;    Score 15.4;  DB 4;  Length 2430;
  Best Local Similarity  94.1%;   Pred. No. 1.1e+02;
  Matches  16;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy      4 TTATCCAATTCACACA 20
          ||||||||||||| ||||
Db   1819 TTATCCAATTTTACACA 1835


RESULT 12
US-08-845-258-40/c
; Sequence 40, Application US/0845258
; Patent No. 6183976
; GENERAL INFORMATION:
;    APPLICANT: Reed, Steven G.
;    APPLICANT: Lodes, Michael J.
;    APPLICANT: Houghton, Raymond
;    APPLICANT: Sleath, Paul R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;    TITLE OF INVENTION:  AND TREATMENT OF B. MICROTI INFECTION
;    NUMBER OF SEQUENCES: 53
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: SEED AND BERRY
;       STREET: 6300 Columbia Center, 701 Fifth Avenue
;       CITY: Seattle
;       STATE: Washington
;       COUNTRY: USA
;       ZIP: 98104
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/845,258
;       FILING DATE: 24-APR-1997
;       CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Maki, David J.
;       REGISTRATION NUMBER: 31,392
;       REFERENCE/DOCKET NUMBER: 210121.426C1
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (206) 622-4900
;       TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 40:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 2430 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
US-08-845-258-40

  Query Match      73.3%;    Score 15.4;  DB 4;  Length 2430;
  Best Local Similarity  94.1%;   Pred. No. 1.1e+02;
  Matches  16;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy      4 TTATCCAATTCACACA 20
          ||||||||||||| ||||
Db    612 TTATCCAATTTTACACA 596


RESULT 13
US-08-990-571-3
; Sequence 3, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
;    APPLICANT: Reed, Steven G. et al.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B. M
;    NUMBER OF SEQUENCES: 79
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: SEED AND BERRY
;       STREET: 6300 Columbia Center, 701 Fifth Avenue
;       CITY: Seattle
;       STATE: Washington
;       COUNTRY: USA
;       ZIP: 98104
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/990,571
;       FILING DATE: 11-DEC-1997
;       CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Maki, David J.
;       REGISTRATION NUMBER: 31,392
;       REFERENCE/DOCKET NUMBER: 210121.426C2
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (206) 622-4900
;       TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 3:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 2430 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
US-08-990-571-3

  Query Match      73.3%;    Score 15.4;  DB 4;  Length 2430;
  Best Local Similarity  94.1%;   Pred. No. 1.1e+02;
  Matches  16;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy      4 TTATCCAATTCACACA 20
          ||||||||||||| ||||
Db   1819 TTATCCAATTTTACACA 1835


RESULT 14
US-08-990-571-40/c
; Sequence 40, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
;    APPLICANT: Reed, Steven G. et al.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B. M
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:   (206) 622-4900
;      TELEFAX:   (206)682-6031
; INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2430 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-723-142A-3

Query Match             73.3%;  Score 15.4;  DB 4;  Length 2430;
Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
Matches   16;  Conservative  0;  Mismatches  1;  Indels   0;  Gaps   0;

Qy         4 TTATCCAATTCACACA 20
             ||||||||||| |||||
Db      1819 TTATCCAATTTACACA 1835

Search completed: June  9, 2003, 11:07:40
Job time : 35.5 secs
```

```
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
  ADDRESSEE:  SEED AND BERRY
  STREET:  6300 Columbia Center, 701 Fifth Avenue
  CITY:  Seattle
  STATE:  Washington
  COUNTRY:  USA
  ZIP:  98104
COMPUTER READABLE FORM:
  MEDIUM TYPE:  Floppy disk
  COMPUTER:  IBM PC compatible
  OPERATING SYSTEM:  PC-DOS/MS-DOS
  SOFTWARE:  PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  US/08/990,571
  FILING DATE:  11-DEC-1997
  CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
  NAME:  Maki, David J.
  REGISTRATION NUMBER:  31,392
  REFERENCE/DOCKET NUMBER:  210121.426C2
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  (206) 622-4900
  TELEFAX:  (206)682-6031
SEQUENCE CHARACTERISTICS:
  LENGTH:  2430 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  single
  TOPOLOGY:  linear
US-08-990-571-40

Query Match             73.3%;  Score 15.4;  DB 4;  Length 2430;
Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
Matches   16;  Conservative  0;  Mismatches  1;  Indels   0;  Gaps   0;

Qy         4 TTATCCAATTCACACA 20
             ||||||||||| |||||
Db       612 TTATCCAATTTACACA 596


RESULT 15
US-08-723-142A-3
; Sequence 3, Application US/08723142A
; Patent No. 6306396
; GENERAL INFORMATION:
;   APPLICANT:  Reed, Steven G.
;   APPLICANT:  Lodes, Michael J.
;   APPLICANT:  Houghton, Raymond
;   APPLICANT:  Sleath, Paul R.
;   TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;   TITLE OF INVENTION:  AND TREATMENT OF B. MICROTI INFECTION
;   NUMBER OF SEQUENCES:  49
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  SEED AND BERRY
;     STREET:  6300 Columbia Center, 701 Fifth Avenue
;     CITY:  Seattle
;     STATE:  Washington
;     COUNTRY:  USA
;     ZIP:  98104
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/723,142A
;     FILING DATE:  01-OCT-1996
;     CLASSIFICATION:  536
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Maki, David J.
;     REGISTRATION NUMBER:  31,392
;     REFERENCE/DOCKET NUMBER:  210121.426
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 09:12:21 ; Search time 59.5 Seconds
(without alignments)
493.954 Million cell updates/sec

Title: US-09-550-163-1_COPY_80_100
Perfect score: 21
Sequence: 1 acttatccaattcacacag 21

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	100.0	372	10	US-09-864-761-33139	Sequence 33139, A
	2	21	100.0	450	10	US-09-864-761-3463	Sequence 3463, Ap
	3	21	100.0	471	10	US-09-864-761-16671	Sequence 16671, A
	4	21	100.0	732	9	US-10-000-151B-5	Sequence 5, Appli
	5	21	100.0	113604	9	US-10-227-195A-1	Sequence 1, Appli
	6	21	100.0	113604	9	US-10-227-195A-2	Sequence 2, Appli
	7	16.4	78.1	339	10	US-09-867-701-9715	Sequence 9715, Ap
c	8	16.4	78.1	461	9	US-09-918-995-31400	Sequence 31400, A
c	9	16.4	78.1	752	9	US-09-776-724A-22	Sequence 22, Appl
c	10	16.4	78.1	893	9	US-10-198-846-2479	Sequence 2479, Ap
c	11	16.4	78.1	6402	9	US-10-187-319-36	Sequence 36, Appl
	12	16.4	78.1	368004	10	US-09-949-654-3	Sequence 3, Appli
	13	16.4	78.1	1830121	9	US-10-329-960-1	Sequence 1, Appli
	14	16.2	77.1	431	9	US-09-918-995-7825	Sequence 7825, Ap
c	15	16.2	77.1	870	10	US-09-969-708-530	Sequence 530, App
c	16	16.2	77.1	1273	9	US-10-091-438-105	Sequence 105, App
c	17	16.2	77.1	1273	10	US-09-764-853-409	Sequence 409, App
c	18	16.2	77.1	1273	10	US-09-764-877-1010	Sequence 1010, Ap
	19	16.2	77.1	1825	9	US-10-198-846-13864	Sequence 13864, A
	20	16.2	77.1	4656	10	US-09-880-107-2268	Sequence 2268, Ap
	21	16	76.2	626	9	US-09-198-846-3625	Sequence 3625, Ap
c	22	15.8	75.2	1013	9	US-09-938-842A-2708	Sequence 2708, Ap
	23	15.8	75.2	1488	10	US-09-755-017-3	Sequence 3, Appli
c	24	15.8	75.2	1763	9	US-10-050-704-76	Sequence 76, Appl
	25	15.8	75.2	4160	9	US-09-822-846-171	Sequence 171, App
c	26	15.8	75.2	5271	12	US-10-044-090-354	Sequence 354, App
	27	15.8	75.2	143306	10	US-09-729-920-3	Sequence 3, Appli
	28	15.4	73.3	307	9	US-09-842-552-18	Sequence 18, Appl
	29	15.4	73.3	447	9	US-10-198-846-5995	Sequence 5995, Ap
	30	15.4	73.3	488	9	US-10-198-846-1261	Sequence 1261, Ap
	31	15.4	73.3	493	10	US-09-878-574-4840	Sequence 4840, Ap
	32	15.4	73.3	495	9	US-10-198-846-209	Sequence 209, App
	33	15.4	73.3	660	9	US-10-198-846-2390	Sequence 2390, Ap
	34	15.4	73.3	667	9	US-10-198-846-7369	Sequence 7369, Ap
	35	15.4	73.3	694	9	US-10-198-846-5280	Sequence 5280, Ap
	36	15.4	73.3	799	9	US-09-764-891-9021	Sequence 9021, Ap
	37	15.4	73.3	838	9	US-10-198-846-2028	Sequence 2028, Ap
	38	15.4	73.3	858	9	US-10-198-846-3321	Sequence 3321, Ap
	39	15.4	73.3	872	9	US-10-198-846-5021	Sequence 5021, Ap
	40	15.4	73.3	882	9	US-10-198-846-9562	Sequence 9562, Ap
	41	15.4	73.3	897	9	US-10-198-846-4454	Sequence 4454, Ap
	42	15.4	73.3	902	9	US-10-198-846-3673	Sequence 3673, Ap
	43	15.4	73.3	912	9	US-10-198-846-2887	Sequence 2887, Ap
	44	15.4	73.3	912	9	US-10-198-846-5534	Sequence 5534, Ap
	45	15.4	73.3	931	9	US-10-198-846-7336	Sequence 7336, Ap

ALIGNMENTS

RESULT 1
US-09-864-761-33139
; Sequence 33139, Application US/09864761
; Patent No. US2002004876341
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR APPLICATION FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR APPLICATION FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR APPLICATION FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR APPLICATION FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR APPLICATION FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR APPLICATION FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR APPLICATION FILING DATE: 2001-01-30

```
;   PRIOR APPLICATION NUMBER: PCT/US01/00670
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/234,687
;   PRIOR FILING DATE: 2000-09-21
;   PRIOR APPLICATION NUMBER: US 09/608,408
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: US 09/774,203
;   PRIOR FILING DATE: 2001-01-29
;   NUMBER OF SEQ ID NOS: 49117
;   SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;   SEQ ID NO 33139
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000120.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
;   OTHER INFORMATION: EST HUMAN HIT: AI654552.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q9Y6J6, EVALUE 8.00e-67
;   OTHER INFORMATION: NT HIT: gi|11526220, EVALUE 0.00e+00
US-09-864-761-33139

Query Match          100.0%;   Score 21;  DB 10;   Length 372;
Best Local Similarity 100.0%;   Pred. No. 3.2;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ACTTTATCCAATTCACACAG 21
         |||||||||||||||||||||
Db     7 ACTTTATCCAATTCACACAG 27

RESULT 2
US-09-864-761-3463
; Sequence 3463, Application US09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;   APPLICANT: Penn, Sharron G.
;   APPLICANT: Rank, David R.
;   APPLICANT: Hanzel, David K.
;   APPLICANT: Chen, Wensheng
;   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;   TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;   FILE REFERENCE: Aeomica-X-1
;   CURRENT APPLICATION NUMBER: US/09/864,761
;   CURRENT FILING DATE: 2001-05-23
;   PRIOR APPLICATION NUMBER: US 60/180,312
;   PRIOR FILING DATE: 2000-02-04
;   PRIOR APPLICATION NUMBER: US 60/207,456
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: US 09/632,366
;   PRIOR FILING DATE: 2000-08-03
;   PRIOR APPLICATION NUMBER: GB 24263.6
;   PRIOR FILING DATE: 2000-10-04
;   PRIOR APPLICATION NUMBER: US 60/236,359
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: PCT/US01/00666
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00669
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00667
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00665
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00668
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00664
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00663
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00662
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00661

;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00670
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/234,687
;   PRIOR FILING DATE: 2000-09-21
;   PRIOR APPLICATION NUMBER: US 09/608,408
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: US 09/774,203
;   PRIOR FILING DATE: 2001-01-29
;   NUMBER OF SEQ ID NOS: 49117
;   SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;   SEQ ID NO 3463
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000052.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.88
US-09-864-761-3463

Query Match          100.0%;   Score 21;  DB 10;   Length 450;
Best Local Similarity 100.0%;   Pred. No. 3.3;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ACTTTATCCAATTCACACAG 21
         |||||||||||||||||||||
Db   162 ACTTTATCCAATTCACACAG 182

RESULT 3
US-09-864-761-16671
; Sequence 16671, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;   APPLICANT: Penn, Sharron G.
;   APPLICANT: Rank, David R.
;   APPLICANT: Hanzel, David K.
;   APPLICANT: Chen, Wensheng
;   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;   TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;   FILE REFERENCE: Aeomica-X-1
;   CURRENT APPLICATION NUMBER: US/09/864,761
;   CURRENT FILING DATE: 2001-05-23
;   PRIOR APPLICATION NUMBER: US 60/180,312
;   PRIOR FILING DATE: 2000-02-04
;   PRIOR APPLICATION NUMBER: US 60/207,456
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: US 09/632,366
;   PRIOR FILING DATE: 2000-08-03
;   PRIOR APPLICATION NUMBER: GB 24263.6
;   PRIOR FILING DATE: 2000-10-04
;   PRIOR APPLICATION NUMBER: US 60/236,359
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: PCT/US01/00666
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00667
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/207,456
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 09/632,366
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/236,359
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00666
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00667
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00664
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00669
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00665
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```
;       PRIOR FILING DATE: 2001-01-30
;       PRIOR APPLICATION NUMBER: PCT/US01/00663
;       PRIOR FILING DATE: 2001-01-30
;       PRIOR APPLICATION NUMBER: PCT/US01/00662
;       PRIOR FILING DATE: 2001-01-30
;       PRIOR APPLICATION NUMBER: PCT/US01/00661
;       PRIOR FILING DATE: 2001-01-30
;       PRIOR APPLICATION NUMBER: PCT/US01/00670
;       PRIOR FILING DATE: 2001-01-30
;       PRIOR APPLICATION NUMBER: US 60/234,687
;       PRIOR FILING DATE: 2000-09-21
;       PRIOR APPLICATION NUMBER: US 09/608,408
;       PRIOR FILING DATE: 2000-06-30
;       PRIOR APPLICATION NUMBER: US 09/774,203
;       PRIOR FILING DATE: 2001-01-29
;       NUMBER OF SEQ ID NOS: 49117
;       SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;       SEQ ID NO 16671
;       LENGTH: 471
;       TYPE: DNA
;       ORGANISM: Homo sapiens
;       FEATURE:
;       OTHER INFORMATION: MAP TO AP000120.1
;       OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.98
;       OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67
US-09-864-761-16671

  Query Match             100.0%;  Score 21;  DB 10;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db    245 ACTTTATCCAATTTCACACAG 265

RESULT 4
US-10-000-151B-5
; Sequence 5, Application US/100001151B
; Publication No. US2003001136A1
; GENERAL INFORMATION:
;       APPLICANT: Balser, Jeffrey R.
;       APPLICANT: George, Alfred L.
;       TITLE OF INVENTION: HUMAN KCR1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
;       FILE REFERENCE: Vanderbilt Ref No. US2003001136A1 VU0120; Attorney Docket No. US2003
;       CURRENT APPLICATION NUMBER: US/10/000,151B
;       CURRENT FILING DATE: 2000-10-30
;       NUMBER OF SEQ ID NOS: 5
;       SOFTWARE: PatentIn version 3.1
;       SEQ ID NO 5
;       LENGTH: 732
;       TYPE: DNA
;       ORGANISM: Homo sapiens
US-10-000-151B-5

  Query Match             100.0%;  Score 21;  DB 9;  Length 732;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db     80 ACTTTATCCAATTTCACACAG 100

RESULT 5
US-10-227-195A-1
; Sequence 1, Application US/10227195A
; Publication No. US2003007633A1
; GENERAL INFORMATION:
;       APPLICANT: Cox, David
;       APPLICANT: Arnold, Deana
;       TITLE OF INVENTION: Haplotype structure of chromosome 21
;       FILE REFERENCE: 1030U1
;       CURRENT APPLICATION NUMBER: US/10/227,195A
;       CURRENT FILING DATE: 2002-11-18
;       NUMBER OF SEQ ID NOS: 2
;       SOFTWARE: FastSEQ for Windows Version 4.0
;       SEQ ID NO 1
;       LENGTH: 113604
;       TYPE: DNA
;       ORGANISM: Human
;       FEATURE:
;       NAME/KEY: misc_feature
;       LOCATION: 7175, 7204, 36973, 66372, 76921, 81512, 88727
;       OTHER INFORMATION: n = G or C
US-10-227-195A-1

  Query Match             100.0%;  Score 21;  DB 9;  Length 113604;
  Best Local Similarity   100.0%;  Pred. No. 8.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db  17482 ACTTTATCCAATTTCACACAG 17502

RESULT 6
US-10-227-195A-2
; Sequence 2, Application US/10227195A
; Publication No. US2003007633A1
; GENERAL INFORMATION:
;       APPLICANT: Cox, David
;       APPLICANT: Arnold, Deana
;       TITLE OF INVENTION: Haplotype structure of chromosome 21
;       FILE REFERENCE: 1030U1
;       CURRENT APPLICATION NUMBER: US/10/227,195A
;       CURRENT FILING DATE: 2002-11-18
;       NUMBER OF SEQ ID NOS: 2
;       SOFTWARE: FastSEQ for Windows Version 4.0
;       SEQ ID NO 2
;       LENGTH: 113604
;       TYPE: DNA
;       ORGANISM: Human
US-10-227-195A-2

  Query Match             100.0%;  Score 21;  DB 9;  Length 113604;
  Best Local Similarity   100.0%;  Pred. No. 8.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db  17482 ACTTTATCCAATTTCACACAG 17502

RESULT 7
US-09-867-701-9715
; Sequence 9715, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
;       APPLICANT: Aglate, Paul A.
;       APPLICANT: Jones, Robert
;       APPLICANT: Harlocker, Susan L.
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;       TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
;       FILE REFERENCE: 210121.497
;       CURRENT APPLICATION NUMBER: US/09/867,701
;       CURRENT FILING DATE: 2001-05-29
;       NUMBER OF SEQ ID NOS: 10912
;       SOFTWARE: FastSEQ for Windows Version 4.0
;       SEQ ID NO 9715
;       LENGTH: 339
;       TYPE: DNA
;       ORGANISM: Homo sapien
US-09-867-701-9715

  Query Match              78.1%;  Score 16.4;  DB 10;  Length 339;
```

Best Local Similarity 94.4%; Pred. No. 3.7e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;

Qy   1 ACTTTATCCAATTCACA 18
          ||||| ||||||||||
Db   160 ACTTTTTCCAATTCACA 177

RESULT 8
US-09-918-995-31400/c
; Sequence 31400, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 31400
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(461)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31400

Query Match 78.1%; Score 16.4; DB 9; Length 461;
Best Local Similarity 94.4%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;

Qy   4 TTATCCAATTCACACAG 21
          ||||| |||||||||
Db   389 TTATCCAAGTTCACACAG 372

RESULT 9
US-09-776-724A-22
; Sequence 22, Application US/09776724A
; Publication No. US2003050455A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 64 Human Secreted Proteins
; FILE REFERENCE: P2011
; CURRENT APPLICATION NUMBER: US/09/776,724A
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/180,909
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 09/669,688
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/229,982
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: PCT/US98/14613
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/052,661
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,872
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,871
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,874
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,873
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,870
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,875

; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/053,440
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,441
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,442
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/056,359
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,725
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,985
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,952
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,989
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,361
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,726
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,724
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,946
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,683
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-776-724A-22

Query Match 78.1%; Score 16.4; DB 9; Length 752;
Best Local Similarity 94.4%; Pred. No. 4.3e+02; Indels 1; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;

Qy   1 ACTTTATCCAATTCACA 18
          ||||||||| |||||||
Db   116 ACTTTATCCCATTTCACA 133

RESULT 10
US-10-198-846-2479/c
; Sequence 2479, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2479
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 456, 466, 486, 532, 537, 546, 579, 580, 588, 605, 631, 699,
; LOCATION: 700, 707, 709, 710, 723, 726, 742, 746, 754, 792, 804, 805,
; LOCATION: 808, 811, 813, 814, 817, 820, 825, 828, 831, 835, 838, 844,
; LOCATION: 846, 854, 857, 858, 862, 866, 873, 879, 888, 892

```
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2479

Query Match        78.1%;    Score 16.4;   DB 9;   Length 893;
Best Local Similarity  94.4%;   Pred. No. 4.4e+02;
Matches   17;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy      4  TTATCCAATTCACACAG 21
           ||| ||||||||||||||
Db    576  TTAACCAATTCACACAG 559

RESULT 11
US-10-187-319-36/c
; Sequence 36, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
;   APPLICANT: Lollar, John S.
;   TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;   NUMBER OF SEQUENCES: 40
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Greenlee, Winner and Sullivan, P. C.
;     STREET: 5370 Manhattan Circle Suite 201
;     CITY: Boulder
;     STATE: Colorado
;     COUNTRY: USA
;     ZIP: 80303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/187,319
;     FILING DATE: 27-Aug-2002
;     CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 09/523,656
;     FILING DATE: 2000-03-10
;     APPLICATION NUMBER: US 09/037,601
;     FILING DATE: 1998-03-10
;     APPLICATION NUMBER: WO PCT/US97/11155
;     FILING DATE: 1997-06-26
;     APPLICATION NUMBER: US 08/670,707
;     FILING DATE: 1996-06-26
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Greenlee, Lorance L.
;     REGISTRATION NUMBER: 27,894
;     REFERENCE/DOCKET NUMBER: 75-95K
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 303/499-8080
;     TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6402 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: <Unknown>
;   MOLECULE TYPE: cDNA to mRNA
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;     ORGANISM: Pig
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..6402
;   SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-10-187-319-36

Query Match        78.1%;    Score 16.4;   DB 9;   Length 6402;
Best Local Similarity  94.4%;   Pred. No. 6.2e+02;
Matches   17;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy      3  TTTATCCAATTCACACA 20
Db    3116  TTTATCCAATTGCACACA 3099

RESULT 12
US-09-949-654-3
; Sequence 3, Application US/09949654
; Patent No. US20020127644A1
; GENERAL INFORMATION:
;   APPLICANT: YAN, Chunhua et al.
;   TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;   TITLE OF INVENTION: AND USES THEREOF
;   FILE REFERENCE: CL000817
;   CURRENT APPLICATION NUMBER: US/09/949,654
;   CURRENT FILING DATE: 2001-09-12
;   PRIOR APPLICATION NUMBER: 60/231,572
;   PRIOR FILING DATE: 2000-09-11
;   NUMBER OF SEQ ID NOS: 5
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 368004
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)..(368004)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-654-3

Query Match        78.1%;    Score 16.4;   DB 10;   Length 368004;
Best Local Similarity  94.4%;   Pred. No. 1.1e+03;
Matches   17;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy      4  TTATCCAATTCACACAG 21
Db  175073  TTATCCAATTGCACACAG 175090

RESULT 13
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US2003009927A1
; GENERAL INFORMATION:
;   APPLICANT: Fleischmann et al.
;   TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra
;   TITLE OF INVENTION: Thereof, and Uses Thereof
;   FILE REFERENCE: PB186P1
;   CURRENT APPLICATION NUMBER: US/10/329,960
;   CURRENT FILING DATE: 2003-01-02
;   PRIOR APPLICATION NUMBER: US 09/643,990
;   PRIOR FILING DATE: 2000-08-23
;   PRIOR APPLICATION NUMBER: US 08/487,429
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: US 08/426,787
;   PRIOR FILING DATE: 1995-04-21
;   NUMBER OF SEQ ID NOS: 1
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (10150)..(10150)
```

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (53369)..(53369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

```
;     NAME/KEY: misc_feature
;     LOCATION: (145171)..(145171)
;     OTHER INFORMATION: n equals a, t, g or c
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (145942)..(145942)
;     OTHER INFORMATION: n equals a, t, g or c
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (147197)..(147197)
;     OTHER INFORMATION: n equals a, t, g or c
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (150841)..(150841)
;     OTHER INFORMATION: n equals a, t, g or c
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (152500)..(152500)
;     OTHER INFORMATION: n equals a, t, g or c
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (152530)..(152530)

Query Match             78.1%;   Score 16.4;  DB 9;  Length 1830121;
Best Local Similarity   94.4%;   Pred. No. 6.8e+02;
Matches  17;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy       1 ACTTTATCCAATTTCACA 18
           ||||||||||||||||||
Db  161151 ACTTTATCCAAGTTCACA 161168


RESULT 14
US-09-918-995-7825
; Sequence 7825, Application US/09918995
; Publication No. US2003007362A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7825
;     LENGTH: 431
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-09-918-995-7825

Query Match             77.1%;   Score 16.2;  DB 9;  Length 431;
Best Local Similarity   85.7%;   Pred. No. 4.7e+02;
Matches  18;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 ACTTTATCCAATTTCACACAG 21
           || |||  ||||||||||||||
Db      83 ACCTGATCAAATTTCACACAG 103


RESULT 15
US-09-969-708-530
; Sequence 530, Application US/09969708
; Patent No. US2002010253A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
```

```
;     PRIOR APPLICATION NUMBER: US/60/237,606
;     PRIOR FILING DATE: 2000-10-03
;     PRIOR APPLICATION NUMBER: US/60/237,608
;     PRIOR FILING DATE: 2000-10-03
;     PRIOR APPLICATION NUMBER: US/60/237,425
;     PRIOR FILING DATE: 2000-10-03
;     SOFTWARE: PatentIn version 3.0
;     NUMBER OF SEQ ID NOS: 658
; SEQ ID NO 530
;     LENGTH: 870
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-09-969-708-530

Query Match             77.1%;   Score 16.2;  DB 10;  Length 870;
Best Local Similarity   85.7%;   Pred. No. 5.4e+02;
Matches  18;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 ACTTTATCCAATTTCACACAG 21
           || |||||||||  |  |||||
Db     524 ACTTTATCAAATTTGATACAG 544


Search completed: June   9, 2003, 12:10:08
Job time : 124.5 secs
```

```
                    GenCore version 5.1.6
               Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 08:36:50 ; Search time 883 Seconds
                                      (without alignments)
                                      385.170 Million cell updates/sec

Title:          US-09-550-163-1_COPY_80_100
Perfect score:  21
Sequence:       1 acttatccaattcacacag 21

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
                1: em_estba:*
                2: em_esthum:*
                3: em_estin:*
                4: em_estmu:*
                5: em_estov:*
                6: em_estpl:*
                7: em_estro:*
                8: em_htc:*
                9: gb_est1:*
               10: gb_est2:*
               11: gb_htc:*
               12: gb_est3:*
               13: gb_est4:*
               14: gb_est5:*
               15: em_estfun:*
               16: em_estom:*
               17: gb_gss:*
               18: em_gss_hum:*
               19: em_gss_inv:*
               20: em_gss_pln:*
               21: em_gss_vrt:*
               22: em_gss_fun:*
               23: em_gss_mam:*
               24: em_gss_mus:*
               25: em_gss_other:*
               26: em_gss_pro:*
               27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	21	100.0	311	10	AW869303	AW869303 MR3-SN006
c	2	21	100.0	746	12	BG221966	BG221966 RST41783
c	3	21	100.0	803	12	BG208163	BG208163 RST27654
c	4	18.4	87.6	419	12	BE847728	BE847728 uv93d05.y
c	5	18.4	87.6	561	17	AQ785939	AQ785939 HS_3053_A
c	6	18.4	87.6	666	14	BQ506222	BQ506222 EST613637

c	7	18.4	87.6	689	7	BG595346	BG595346 EST494024
c	8	17.8	84.8	191	12	BF092401	BF092401 MR4-TN010
c	9	17.8	84.8	358	12	BG152818	BG152818 nad07h02.
c	10	17.8	84.8	380	17	AQ207086	AQ207086 HS_3235_A
c	11	17.8	84.8	389	13	BI048187	BI048187 MR4-ST024
c	12	17.8	84.8	396	13	AW784520	AW784520 zb64f06.g
c	13	17.8	84.8	427	12	BF091642	BF091642 MR1-TN004
c	14	17.8	84.8	429	17	AQ037472	AQ037472 CIT-HSP-2
c	15	17.8	84.8	432	17	AQ738020	AQ738020 HS_2063_B
c	16	17.8	84.8	480	17	B88293	B88293 CIT-HSP-206
c	17	17.8	84.8	661	17	B28338	B28338 F13P23TF IG
c	18	17.8	84.8	753	17	B47468	B47468 CIT97BSK-A-
c	19	17.8	84.8	1155	13	BM452226	BM452226 AGENCOURT
c	20	17.4	82.9	412	17	AZ039490	AZ039490 RPCI-23-3
c	21	17.4	82.9	532	12	BF198473	BF198473 248413 MA
c	22	17.4	82.9	632	17	AZ595882	AZ595882 1M0408K18
c	23	17.4	82.9	672	17	BH177640	BH177640 010_M_13-
c	24	17.4	82.9	672	17	CN507K51	AL614599 T3 end of
c	25	17	82.9	1201	17	CNS00102	AL054605 Drosophil
c	26	17	81.0	278	14	CNS00102	AL054605 Drosophil
c	27	17	81.0	695	17	AG095529	AG095529 Pan trogl
c	28	17	81.0	746	17	AG170583	AG170583 Pan trogl
c	29	17	81.0	820	17	BF695254	BF695254 602080491
c	30	16.8	80.0	217	10	BB419925	BB419925 BB419925
c	31	16.8	80.0	241	10	BB082597	BB082597 BB082597
c	32	16.8	80.0	259	10	AV326138	AV326138 AV326138
c	33	16.8	80.0	261	10	AV326162	AV326162 AV326162
c	34	16.8	80.0	261	10	AV326205	AV326205 AV326205
c	35	16.8	80.0	265	17	AQ107464	AQ107464 HS_3108_A
c	36	16.8	80.0	269	9	AA327900	AA327900 EST31302
c	37	16.8	80.0	272	10	BB486188	BB486188 BB486188
c	38	16.8	80.0	280	10	BB302773	BB302773 BB302773
c	39	16.8	80.0	282	10	BB081516	BB081516 BB081516
c	40	16.8	80.0	282	10	BB555197	BB555197 BB555197
c	41	16.8	80.0	283	14	Z20463	Z20463 HSAAACEFR T
c	42	16.8	80.0	289	10	BB366709	BB366709 BB366709
c	43	16.8	80.0	293	10	BB348981	BB348981 BB348981
c	44	16.8	80.0	297	10	BB086469	BB086469 BB086469
c	45	16.8	80.0	298	10	BB044070	BB044070 BB044070

ALIGNMENTS

```
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/script1/gethtml2.pl?i=&t2=MR3-SN0067-240
            400-006-f11&t3=2000-04-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 74.

FEATURES
     source          1..311
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="SN0067"
                     /dev_stage="Adult"
                     /note="organ: stomach normal; Vector: puc18; Site_1: SmaI;
                     Site 2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

BASE COUNT     60 a     80 c     76 g     95 t
ORIGIN

Query Match              100.0%;     Score 21;    DB 10;    Length 311;
Best Local Similarity    100.0%;     Pred. No. 38;
Matches    21;   Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db  176 ACTTTATCCAATTTCACACAG 156

RESULT 2
BG221966
LOCUS       BG221966                746 bp    mRNA      linear    EST 21-APR-2001
DEFINITION  RST41783 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG221966
VERSION     BG221966.1  GI:13747987
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 746)
AUTHORS     Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,J., Danzig,J., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
TITLE       Creation of genome-wide protein expression libraries using random
            activation of gene expression
JOURNAL     Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE     21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 547.
FEATURES
     source          1..746
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression'
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
```

```
BASE COUNT    241 a    138 c    123 g    243 t    1 others
ORIGIN

Query Match              100.0%;     Score 21;    DB 12;    Length 746;
Best Local Similarity    100.0%;     Pred. No. 46;
Matches    21;   Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db  532 ACTTTATCCAATTTCACACAG 552

RESULT 3
BG208163/c
LOCUS       BG208163                803 bp    mRNA      linear    EST 21-APR-2001
DEFINITION  RST27654 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG208163
VERSION     BG208163.1  GI:13729850
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 803)
AUTHORS     Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
TITLE       Creation of genome-wide protein expression libraries using random
            activation of gene expression
JOURNAL     Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE     21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 550.
FEATURES
     source          1..803
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression'
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."

BASE COUNT    222 a    172 c    171 g    238 t
ORIGIN

Query Match              100.0%;     Score 21;    DB 12;    Length 803;
Best Local Similarity    100.0%;     Pred. No. 46;
Matches    21;   Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACTTTATCCAATTTCACACAG 21
          |||||||||||||||||||||
Db  682 ACTTTATCCAATTTCACACAG 662

RESULT 4
BE847728/c
LOCUS       BE847728                419 bp    mRNA      linear    EST 26-SEP-2000
DEFINITION  uv93d05.y1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:314729
            5', mRNA sequence.
ACCESSION   BE847728
VERSION     BE847728.1  GI:10306067
KEYWORDS    EST.
SOURCE      house mouse.
```

```
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 419)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Other_ESTs: uv93d05.x1
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:1090541
           Seq primer: -40RP from Gibco.
FEATURES   Location/Qualifiers
   source  1..419
           /organism="Mus musculus"
           /strain="C57BL/6J"
           /db_xref="taxon:10090"
           /clone="IMAGE:3414729"
           /clone_lib="Soares mouse 3NbMS"
           /sex="male"
           /tissue_type="Spleen"
           /dev_stage="4 weeks"
           /lab_host="DH10B"
           /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
           polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
           was primed with a Not I - oligo(dT) primer [5'
           TGTTACCAATCTGAAGTGGGACGGCGGCGCGTTTTTTTTTTTTTTTTTTTT
           3']; double-stranded cDNA was ligated to Eco RI adaptors
           (Pharmacia), digested with Not I and cloned into the Not I
           and Eco RI sites of the modified pT7T3 vector. RNA
           provided by Dr. Bertrand Jordan. Library went through
           three rounds of normalization, and was constructed by
           Bento Soares and M.Fatima Bonaldo."
BASE COUNT 136 a   62 c   90 g   131 t
ORIGIN

Query Match              87.6%;  Score 18.4;  DB 12;  Length 419;
  Best Local Similarity  95.0%;  Pred. No. 5.5e+02;
  Matches  19;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          1 ACTTTATCCAATTTCACACA 20
              |||||| |||||||||||||
Db        217 ACTTTTTCCAATTTCACACA 198

RESULT 5
AQ785939
LOCUS       AQ785939              561 bp    DNA      linear   GSS 03-AUG-1999
DEFINITION  HS_3053_A2_G02_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3053 Col=4 Row=M, DNA sequence.
ACCESSION   AQ785939
VERSION     AQ785939.1  GI:5693563
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 561)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3053 row: M column: 4
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 561.
FEATURES    Location/Qualifiers
   source   1..561
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="Plate=3053 Col=4 Row=M"
            /clone_lib="CIT Approved Human Genomic Sperm Library D"
            /sex="male"
            /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
            E-Coli DH10B"
BASE COUNT  181 a   92 c   72 g   211 t    5 others
ORIGIN

Query Match              87.6%;  Score 18.4;  DB 17;  Length 561;
  Best Local Similarity  95.0%;  Pred. No. 5.9e+02;
  Matches  19;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          2 CTTTATCCAATTCACACAG 21
              ||| ||||||||||||||
Db        472 CTTTATCCAATTCCACACAG 491

RESULT 6
BQ506222
LOCUS       BQ506222               666 bp    mRNA     linear   EST 22-JUL-2002
DEFINITION  EST613637 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMG192
            5' end, mRNA sequence.
ACCESSION   BQ506222
VERSION     BQ506222.2  GI:21922104
KEYWORDS    EST.
SOURCE      potato.
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1 (bases 1 to 666)
AUTHORS     Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., and
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
TITLE       Generation of a set of potato cDNA clones for microarray analyses
JOURNAL     Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21365091.
            Other ESTs: EST613638
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T3.
FEATURES    Location/Qualifiers
   source   1..666
            /organism="Solanum tuberosum"
            /cultivar="Kennebec or Binjte"
            /db_xref="taxon:4113"
            /clone="STMG192"
            /clone_lib="Generation of a set of potato cDNA clones for
            microarray analyses mixed potato tissues"
            /tissue_type="mixed tissues"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; supplier: Combination of untreated and Phytophthora
            infestans-treated libraries of stolons, leaves, leaflets,
            axillary buds of stem explants, petioles, germinating eyes
```

```
                              , tubers, or roots."
BASE COUNT      200 a    159 c    124 g    183 t
ORIGIN

Query Match              87.6%;  Score 18.4;  DB 14;  Length 666;
Best Local Similarity    95.0%;  Pred. No. 6.1e+02;
Matches    19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1 ACTTTATCCAATTCACACA 20
         |||||||||||||||||||
Db   229 ACTTTATCCAATTCACACA 248


RESULT 7
BG595346
LOCUS       BG595346      689 bp     mRNA    linear   EST 12-APR-2001
DEFINITION  EST494024 cSTS Solanum tuberosum cDNA clone cSTS1019 5' sequence,
            mRNA sequence.
ACCESSION   BG595346
VERSION     BG595346.1  GI:13613486
KEYWORDS    EST.
SOURCE      potato.
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 689)
AUTHORS     van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE       Generations of ESTs from sprouting potato eyes
JOURNAL     Unpublished (2000)
COMMENT     Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1..689
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTS1019"
                     /clone_lib="cSTS"
                     /tissue_type="sprouting eyes from tubers"
                     /dev_stage="12-14 weeks post harvest"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Various sizes of nprouting eyes (72mm to 15mm) were
                     taken from tubers. The tubers were incubated at 26C in the
                     dark for 2-3 weeks prior to sprouting. The eyes were
                     frozen in liquid nitrogen immediately upon removal from
                     tubers."
BASE COUNT      209 a    164 c    128 g    188 t
ORIGIN

Query Match              87.6%;  Score 18.4;  DB 12;  Length 689;
Best Local Similarity    95.0%;  Pred. No. 6.1e+02;
Matches    19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1 ACTTTATCCAATTCACACA 20
         |||||||||||||||||||
Db   229 ACTTTATCCAATTCACACA 248


RESULT 8
BF092401/c
LOCUS       BF092401      191 bp     mRNA    linear   EST 19-OCT-2000
DEFINITION  MR4-TN0109-120900-202-f11 TN0109 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF092401
VERSION     BF092401.1  GI:10898215
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 191)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
            M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tcl=&t2=MR4-TN0109-120
            900-202-f11&t3=2000-09-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 2
            High quality sequence stop: 190.
FEATURES             Location/Qualifiers
     source          1..191
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="TN0109"
                     /dev_stage="Adult"
                     /note="Organ: testis normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       55 a     47 c     42 g     47 t
ORIGIN

Query Match              84.8%;  Score 17.8;  DB 12;  Length 191;
Best Local Similarity    90.5%;  Pred. No. 8.6e+02;
Matches    19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     1 ACTTTATCCAATTCACACAG 21
         ||||||||||||| |||
Db   101 ACTTTATCCAATTGATACAG 81


RESULT 9
BG152818/c
LOCUS       BG152818      358 bp     mRNA    linear   EST 05-FEB-2001
DEFINITION  nad07h02.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3432819 3'
            similar to SW:RS4Y_HUMAN P22090 40S RIBOSOMAL PROTEIN S4, Y
            ISOFORM.; mRNA sequence.
ACCESSION   BG152818
VERSION     BG152818.1  GI:12664848
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck.M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:.
info@image.llnl.gov
Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
     source          Location/Qualifiers
                     1..358
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3432819"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT       96 a      93 c      72 g      97 t
ORIGIN

Query Match               84.8%;   Score 17.8;  DB 12;   Length 358;
Best Local Similarity     90.5%;   Pred. No. 9.8e+02;
Matches  19;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     1 ACTTTATCCAATTCACACAG 21
           |||||||||||| ||||||
Db   276 ACTTTATCCAATTTGATACAG 256

RESULT 10
AQ207086/c
LOCUS       AQ207086    380 bp    DNA    linear    GSS 17-SEP-1998
DEFINITION  HS_3235_A2_B11_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3235 Col=22 Row=C, DNA sequence.
ACCESSION   AQ207086
VERSION     AQ207086.1  GI:3617656
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 380)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3235 row: C column: 22
            Class: BAC ends
            High quality sequence stop: 380.

FEATURES
            Location/Qualifiers

     source
                     1..380
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3235 Col=22 Row=C"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"

BASE COUNT      106 a     103 c      69 g     102 t
ORIGIN

Query Match               84.8%;   Score 17.8;  DB 17;   Length 380;
Best Local Similarity     90.5%;   Pred. No. 9.9e+02;
Matches  19;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     1 ACTTTATCCAATTCACACAG 21
           |||||||||||| ||||||
Db    95 ACTTTATCCAAGGTCACACAG 75

RESULT 11
BI048187/c
LOCUS       BI048187    389 bp    mRNA    linear    EST 14-JUN-2001
DEFINITION  MR4-ST0240-300301-034-d11 ST0240 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI048187
VERSION     BI048187.1  GI:14454809
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20026663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ST0240-
            300301-034-d11&t3=2001-03-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 104
            High quality sequence stop: 206.

FEATURES
     source          Location/Qualifiers
                     1..389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0240"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No.196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

BASE COUNT      101 a      86 c     113 g      89 t
ORIGIN
```

```
Query Match    Best Local Similarity   84.8%;   Score  17.8;  DB 13;   Length 389;
Matches  19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy          1  ACTTATCCAATTCACACAG 21
               |||||||||||| |||||||
Db        274  ACTTTATCCAATTTGATACAG 254

RESULT 12
AW784520/c
LOCUS       AW784520   396 bp   mRNA   linear   EST 22-MAY-2000
DEFINITION  zb64f06.g1 Canis cDNAs from mdck cells Canis familiaris cDNA clone
            zb64f06 5', mRNA sequence.
ACCESSION   AW784520
VERSION     AW784520.1  GI:7838968
KEYWORDS    EST.
SOURCE      dog.
ORGANISM    Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 396)
AUTHORS     McCombie,W.R., See,L.-H., Baker,J.P., Bahret,A., Dedhia,N.N., de la
            Bastide,M., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,
            Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R., Rodriguez,M.A.,
            Shah,R.S., Shekher,M., Spiegel,L.A., Toth,K., Vil,M.D. and Hannon
            ,G.J.
TITLE       Expressed sequence tags from Canis familiaris (dog)
JOURNAL     Unpublished (2000)
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: zb64 row: f column: 06
            Seq primer: -40M13RevUniv
            High quality sequence stop: 396.
FEATURES    Location/Qualifiers
     source 1..396
            /organism="Canis familiaris"
            /db_xref="taxon:9615"
            /clone="zb64f06"
            /clone_lib="Canis cDNAu from mdck cells"
            /note="Vector: Lambda Zap II; The library was provided by
            Greg Hannon (Cold Spring Harbor Laboratory). This
            library is oligo(dT) primed using stratagene zap cDNA
            synthesis kit. It was made from exponentially growing mdck
            cells. Please contact Greg Hannon (hannon@cshl.org) with
            any library related inquiries."
BASE COUNT  71 a  125 c  82 g  118 t
ORIGIN

Query Match    Best Local Similarity   84.8%;   Score  17.8;  DB 10;   Length 396;
Matches  19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;
                                      Pred. No. 1e+03;

Qy          1  ACTTATCCAATTCACACAG 21
               ||||| ||||||| |||||
Db        100  ACTTAATCCAATTTCACAAAG 80

RESULT 13
BF091642/c
LOCUS       BF091642   427 bp   mRNA   linear   EST 19-OCT-2000
DEFINITION  MR1-TN0045-130900-012-a09 TN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF091642
VERSION     BF091642.1  GI:10897352
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 427)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR1-TN0045-130
            900-012-a09&t3=2000-09-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 427.
FEATURES    Location/Qualifiers
     source 1..427
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="TN0045"
            /dev_stage="Adult"
            /note="Organ: testis normal; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the puc 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
BASE COUNT  116 a  105 c  93 g  113 t
ORIGIN

RESULT 14
AQ037472
Query Match    Best Local Similarity   84.8%;   Score  17.8;  DB 12;   Length 427;
Matches  19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;
                                      Pred. No. 1e+03;

Qy          1  ACTTTATCCAATTCACACAG 21
               ||||||||||||| ||||||
Db        338  ACTTTATCCAATTTGATACAG 318

RESULT 14
AQ037472
LOCUS       AQ037472   429 bp   DNA   linear   GSS 11-JUL-1998
DEFINITION  CIT-HSP-2337A2.TF CIT-HSP Homo sapiens genomic clone 2337A2, DNA
            sequence.
ACCESSION   AQ037472
VERSION     AQ037472.1  GI:3303304
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 429)
AUTHORS     Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
JOURNAL     Unpublished (1998)
COMMENT     Other GSSs: CIT-HSP-2337A2.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
```

```
                                              E-Coli DH10B"                      1 others

BASE COUNT      138 a      91 c      74 g      128 t
ORIGIN

Query Match                84.8%;    Score 17.8;   DB 17;    Length 432;
Best Local Similarity      90.5%;    Pred. No. 1e+03;
Matches    19;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ACTTTATCCAATTTCACACAG 21
         ||||| ||||||||||||||
Db    42 ACTTTGCCCAATTTCACACAG 62

Search completed: June  9, 2003, 10:05:49
Job time : 889 secs
```

```
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES          Location/Qualifiers
    source        1..429
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="2337A2"
                  /clone_lib="CIT-HSP"
                  /sex="Male"
                  /cell_type="Sperm"
                  /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                  HindIII"

BASE COUNT     119 a     106 c      78 g      126 t
ORIGIN

Query Match               84.8%;   Score 17.8;   DB 17;   Length 429;
Best Local Similarity     90.5%;   Pred. No. 1e+03;
Matches    19;  Conservative   0;  Mismatches   0;  Indels   2;  Gaps   0;

Qy     1 ACTTTATCCAATTTCACACAG 21
         ||||| ||||||||||
Db    99 ACTTTGTCCAATGTCACACAG 119

RESULT 15
AQ738020
LOCUS       AQ738020               432 bp    DNA       linear   GSS 16-JUL-1999
DEFINITION  HS_2063_B2_B03_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2063 Col=6 Row=D, DNA sequence.
ACCESSION   AQ738020
VERSION     AQ738020.1  GI:515542
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 432)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            BAC end Web Server: http://www.htsc.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            Plate: 2063  row: D  column: 6
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 432.
FEATURES          Location/Qualifiers
    source        1..432
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="Plate=2063 Col=6 Row=D"
                  /clone_lib="CIT Approved Human Genomic Sperm Library D"
                  /sex="male"
                  /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
```

```
                        GenCore version 5.1.6
                        Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  9, 2003, 11:54:30 ; Search time 1051 Seconds
                                     (without alignments)
                                     2789.137 Million cell updates/sec

Title:     US-09-550-163-1_COPY_80_260
Perfect score: 181
Sequence:  1 acttatccaattcacaca.........gattggaatgtctcttca 181

Scoring table: OLIGO_NUC
           Gapop 60.0 , Gapext 60.0
Searched:  1615406 6 seqs, 809774337 6 residues
Word size : 0

Total number of hits satisfying chosen parameters:   32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 1000 summaries

Database :   EST:*
             1: em_estba:*
             2: em_esthum:*
             3: em_estin:*
             4: em_estmu:*
             5: em_estov:*
             6: em_estpl:*
             7: em_estro:*
             8: em_htc:*
             9: gb_est1:*
            10: gb_est2:*
            11: gb_htc:*
            12: gb_est3:*
            13: gb_est4:*
            14: gb_est5:*
            15: em_estfun:*
            16: em_estom:*
            17: gb_gss:*
            18: em_gss_hum:*
            19: em_gss_inv:*
            20: em_gss_pln:*
            21: em_gss_vrt:*
            22: em_gss_fun:*
            23: em_gss_mam:*
            24: em_gss_mus:*
            25: em_gss_other:*
            26: em_gss_pro:*
            27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

c	Result No.	Score	% Match	Query Length	DB	ID	Description
	1	151	83.4	410	9	AI962650	AI962650 wq42e03.x
	2	151	83.4	429	9	AI654552	AI654552 wb48b12.x
	3	143	79.0	372	9	AI246239	AI246239 qi29g04.x
	4	119	65.7	391	9	AI339609	AI339609 qq42a07.x
c	5	117	64.6	803	12	BG208163	BG208163 RST27654
c	6	84	46.4	746	12	BG221966	BG221966 RST41783
c	7	53	29.3	311	10	AW869303	AW869303 MR3-SN006
	8	45	24.9	311	10	AW869303	AW869303 MR3-SN006
	9	35	19.3	121	9	AA935321	AA935321 oo71g09.s
	10	30	16.6	351	13	BG938225	BG938225 1Ab014A12
	11	26	14.4	470	14	D85797	D85797 D85797 Rat
	12	26	14.4	1003	11	BG261965	BG261965 602373784
	13	26	14.4	1691	12	AK008619	AK008619 Mus muscu
	14	22	12.2	314	14	BB564873	BB564873 BB564873
	15	21	11.6	365	14	H77654	H77654 yu24d09.s1
c	16	19	10.5	287	17	BH827709	BH827709 BACPP26-G
c	17	19	10.5	341	17	BH822585	BH822585 BACPP17-F
	18	19	10.5	529	17	BH292512	BH292512 CH230-44E
	19	19	10.5	799	17	BH302149	BH302149 CH230-85A
	20	19	10.5	938	9	AL557074	AL557074 AL557074
	21	18	9.9	139	13	BH644242	BH644242 1008056H0
c	22	18	9.9	169	10	BM416858	BM416858 953004F12
	23	18	9.9	251	10	AV358849	AV358849 AV358849
c	24	18	9.9	414	9	AW921614	AW921614 EST352918
	25	18	9.9	419	9	AA984762	AA984762 am90g02.s
	26	18	9.9	425	9	BE478577	BE478577 162895 BA
c	27	18	9.9	471	12	AL040827	AL040827 DKFZp434F
	28	18	9.9	479	13	BE941069	BE941069 EST420648
	29	18	9.9	580	9	BI742594	BI742594 kx13a05.y
	30	18	9.9		17	AI649434	AI649434 603004D02
c	31	18	9.9		17	AG009962	AG009962 Homo sapi
	32	18	9.9	682	17	AG009955	AG009955 Homo sapi
	33	18	9.9	690	17	AG009958	AG009958 Homo sapi
	34	18	9.9	692	17	AG009957	AG009957 Homo sapi
	35	18	9.9	700	17	AG009959	AG009959 Homo sapi
c	36	18	9.9	702	17	BH556024	BH556024 BOGWT27TF
	37	18	9.9	704	17	AQ362915	AQ362915 nbxb0051A
	38	17	9.4	708	9	AF136410	AF136410 AF136410
	39	17	9.4	714	17	BJ172818	BJ172818 BJ172818
	40	17	9.4	733	13	BH444542	BH444542 BOGIV33TR
	41	17	9.4	736	13	BM072946	BM072946 MEST56-C1
	42	17	9.4	755	17	BH571181	BH571181 BOHKK72TF
	43	17	9.4	770	17	BI656335	BI656335 603282642
	44	17	9.4	794	11	BH738785	BH738785 BOMPT62TR
c	45	17	9.4	1012	17	CNS03ZRF	CNS03ZRF Tetraodon
c	46	17	9.4	2876	17	AY109642	AY109642 Zea mays
c	47	17	9.4	139	17	BH644242	BH644242 1008056H0
c	48	17	9.4	175	10	AW863544	AW863544 MR3-SN000
	49	17	9.4	181	9	BH231068	BH231068 1006160E0
c	50	17	9.4	208	9	AW495284	AW495284 UI-M-BH3-
	51	17	9.4	227	10	AV101033	AV101033 AV101033
c	52	17	9.4	239	17	AI847595	AI847595 UI-M-AP1-
c	53	17	9.4	240	14	AV358518	AV358518 AV358518
	54	17	9.4	261	9	BB263233	BB263233 BB263233
c	55	17	9.4	278	9	BQ371231	BQ371231 RC3-FN015
	56	17	9.4	281	12	AT001562	AT001562 AT001562
	57	17	9.4	300	12	AU116118	AU116118 AU116118
c	58	17	9.4	301	12	BF288220	BF288220 EST452811
	59	17	9.4	303	17	BF415918	BF415918 UI-R-CA0-
	60	17	9.4	309	17	BF400943	BF400943 UI-R-CA0-
c	61	17	9.4	317	17	AZ766217	AZ766217 1M0563B09
c	62	17	9.4	327	10	BB557670	BB557670 BB557670
c	63	17	9.4	331	10	AA801114	AA801114 EST190611
	64	17	9.4	344	10	AV946894	AV946894 AV946894
c	65	17	9.4	355	12	AL127464	AL127464 Fugu rubr
c	66	17	9.4	356	9	BES80642	BES80642 kp79g12.y
c	67	17	9.4	356	17	BF406004	BF406004 UI-R-CA1-
c	68	17	9.4	356	9	AZ881314	AZ881314 RPCI-23-1
	69	17	9.4	371	9	AQ004063	AQ004063 CIT-HSP-2
	70	17	9.4	374	14	AA650533	AA650533 nt05h04.s
	71	17	9.4	375	17	AA650575	AA650575 nt06d07.s
	72	17	9.4	377	12	BQ487193	BQ487193 1091051G0
c	73	17	9.4	384	12	AQ135150	AQ135150 HS_3056_B
c	74	17	9.4	387	9	BF047677	BF047677 dc81c01.y
	75	17	9.4	406	9	AA117808	AA117808 mo65a04.r
	76	17	9.4	411	9	AU090036	AU090036 AU090036
c	77	17	9.4	412	9	AA751686	AA751686 96AS0449
c	78	17	9.4	420	9	AJ485267	AJ485267 AJ485267
c	79	17	9.4	431	12	BF410087	BF410087 UI-R-CA1-

Left block (entries 80–152)

C	No.	Tm	Len	Pos	N	Accession	Name
c	80	9.4	17	431	14	W34878	W34878 mc33f03.r1
c	81	9.4	17	432	13	BM27481	BM275481 PfESToaa5
c	82	9.4	17	442	12	BG398657	BG398657 602440263
c	83	9.4	17	444	17	BH371226	BH371226 AG-ND-142
c	84	9.4	17	447	10	BE051722	BE051722 za87c09.g
	85	9.4	17	454	13	BJ031111	BJ031111 BJ031111
	86	9.4	17	457	13	AW759310	AW759310 s139f06.y
	87	9.4	17	460	10	BE510577	BE510577 946053H06
	88	9.4	17	460	12	BF410138	BF410138 UI-R-CA1-
	89	9.4	17	461	12	BF406119	BF406119 UI-R-CA1-
	90	9.4	17	461	12	BF410303	BF410303 UI-R-CA1-
	91	9.4	17	463	17	AQ691881	AQ691881 HS_5400_B
	92	9.4	17	465	12	BF416536	BF416536 UI-R-CA0-
	93	9.4	17	466	13	BM274777	BM274777 PfESToaa7
	94	9.4	17	471	13	BI441364	BI441364 ic52f12.y
c	95	9.4	17	471	13	BI703912	BI703912 rm07g12.y
	96	9.4	17	472	17	BH026925	BH026925 RPCI-24-3
	97	9.4	17	474	13	BF413079	BF413079 UI-R-CA1-
	98	9.4	17	475	12	BF410450	BF410450 UI-R-CA1-
	99	9.4	17	475	12	BF413743	BF413743 UI-R-CA0-
c	100	9.4	17	476	12	AW212713	AW212713 uo10h02.x
c	101	9.4	17	476	12	BF401417	BF401417 UI-R-CA0-
c	102	9.4	17	476	12	BF413415	BF413415 UI-R-CA0-
c	103	9.4	17	478	13	BM275940	BM275940 PfESToaa8
	104	9.4	17	478	17	BF413078	BF413078 UI-R-CA0-
c	105	9.4	17	481	12	BG555274	BG555274 df01d06.x
	106	9.4	17	482	10	BB701577	BB701577 BB701577
c	107	9.4	17	484	9	AU089979	AU089979 AU089979
c	108	9.4	17	485	14	BM871889	BM871889 mgns015xL
c	109	9.4	17	488	12	BF775260	BF775260 285372_MA
	110	9.4	17	490	9	AU089958	AU089958 AU089958
c	111	9.4	17	492	13	BM274748	BM274748 PfESToaa7
c	112	9.4	17	499	13	BI789832	BI789832 ic44a09.y
c	113	9.4	17	501	13	BM276046	BM276046 PfESToaa6
	114	9.4	17	504	13	BM275867	BM275867 PfESToaa8
c	115	9.4	17	504	13	BM275949	BM275949 PfESToaa8
	116	9.4	17	505	10	BE029193	BE029193 kp26a08.y
	117	9.4	17	511	12	AZ999669	AZ999669 2M0287B11
c	118	9.4	17	512	17	BF406001	BF406001 UI-R-CA1-
c	119	9.4	17	512	12	AZ929244	AZ929244 479.dif23
	120	9.4	17	514	12	BF419694	BF419694 UI-R-CA0-
c	121	9.4	17	514	13	AZ928637	AZ928637 479.dif13
	122	9.4	17	515	17	AZ928491	AZ928491 479.dif10
c	123	9.4	17	518	9	AI958218	AI958218 fc92g12.y
c	124	9.4	17	520	12	AZ929088	AZ929088 479.dif21
c	125	9.4	17	521	13	BE983632	BE983632 UI-M-CG0p
c	126	9.4	17	522	17	AI391177	AI391177 mc33f03.y
c	127	9.4	17	522	10	AW587905	AW587905 kp09f01.y
c	128	9.4	17	532	13	BM275295	BM275295 PfESToaa6
	129	9.4	17	535	13	BM275547	BM275547 PfESToaa5
c	130	9.4	17	536	13	BI678368	BI678368 SWS63_SWS
c	131	9.4	17	537	12	BE989182	BE989182 UI-M-CG0p
c	132	9.4	17	538	17	BH308335	BH308335 CH230-38B
c	133	9.4	17	549	17	AQ757644	AQ757644 HS_2274_B
c	134	9.4	17	553	17	BH537701	BH537701 BOHLT67TF
	135	9.4	17	562	9	AI180401	AI180401 EST224146
c	136	9.4	17	566	12	BF410287	BF410287 UI-R-CA1-
c	137	9.4	17	569	10	AW199521	AW199521 da12h05.x
c	138	9.4	17	576	10	AW782315	AW782315 da87c05.x
	139	9.4	17	580	9	AI649434	AI649434 603004D02
	140	9.4	17	585	12	BG016405	BG016405 df88d11.x
c	141	9.4	17	585	12	BF414554	BF414554 UI-R-CA0-
c	142	9.4	17	586	12	DR3C22T	AL743452 Danio_rer
	143	9.4	17	588	9	AJ436041	AJ436041 AJ436041
c	144	9.4	17	594	17	AZ277521	AZ277521 RPCI-23-1
c	145	9.4	17	596	12	AZ411300	AZ411300 1M0184K17
	146	9.4	17	606	14	BQ211513	BQ211513 UI-R-DY1-
c	147	9.4	17	611	12	BE948717	BE948717 UI-M-BH3-
c	148	9.4	17	616	17	BH655725	BH655725 BOMAK55TR
c	149	9.4	17	616	17	FR0039950	AL127445 imageqc_4
	150	9.4	17	617	13	BM172676	BM172676 Fugu_rubr
	151	9.4	17	627	13	BJ473942	BJ473942 BJ473942
	152	9.4	17	627	14	BQ386431	BQ386431 NISC_mn18

Right block (entries 153–225)

C	No.	Len	Tm	Pos	N	Accession	Name
	153	17	9.4	631	13	BJ089942	BJ089942
c	154	17	9.4	632	17	AZ595882	1M0408K18
c	155	17	9.4	633	13	BJ448109	BJ448109
c	156	17	9.4	634	13	BJ064176	BJ064176
c	157	17	9.4	641	17	AG095094	Pan_trogl
c	158	17	9.4	659	10	BB620211	BB620211
c	159	17	9.4	661	13	BM206425	C0297D07-
	160	17	9.4	662	17	CNS0G51	Tetraodon
	161	17	9.4	663	17	AZ570619	276pVE11
	162	17	9.4	665	9	AL796128	AL796128
	163	17	9.4	665	14	BQ402797	GA_Ed005
	164	17	9.4	666	10	BB638370	BB638370
	165	17	9.4	667	10	BE261578	60114928
	166	17	9.4	670	12	BF071772	db57g11.x
	167	17	9.4	671	10	BB622794	BB622794
c	168	17	9.4	673	17	BH040661	RPCI-24-3
	169	17	9.4	674	14	BQ413226	GA_Ed007
	170	17	9.4	675	17	AQ263171	CITBI-E1-
	171	17	9.4	680	10	AV918922	AV918922
c	172	17	9.4	695	17	AG095529	Pan_trogl
	173	17	9.4	696	13	BI412552	60290859
	174	17	9.4	698	13	BJ475913	BJ475913
	175	17	9.4	707	10	AW766851	de92c06.x
	176	17	9.4	719	12	BF231745	da72d02.x
	177	17	9.4	721	13	AZ084741	RPCI-23-3
	178	17	9.4	722	10	BI650200	60329629
c	179	17	9.4	725	10	AW707294	JAA000631
c	180	17	9.4	727	10	BH062950	RPCI-24-3
	181	17	9.4	730	13	BE303505	60108661
	182	17	9.4	740	10	BJ476439	BJ476439
	183	17	9.4	746	17	AV918005	AV918005
	184	17	9.4	749	13	AG170583	Pan_trogl
	185	17	9.4	755	13	AZ981579	2M0262L03
	186	17	9.4	758	17	BM072946	MEST56-C1
c	187	17	9.4	791	17	CNS03FMK	Tetraodon
	188	17	9.4	820	17	BH737159	BOMFX57TF
	189	17	9.4	874	12	BF695254	BF695254
	190	17	9.4	877	17	BF695254	602080491
	191	17	9.4	885	12	BG480838	602529967
	192	17	9.4	887	12	B12696	F27N8-T7_IG
c	193	17	9.4	898	13	BF232885	602023913
c	194	17	9.4	905	12	BG173194	602335576
	195	17	9.4	937	13	BI078173	602872492
	196	17	9.4	954	17	AL409998	T3_end_of
c	197	17	9.4	968	12	BG114301	602286063
	198	17	9.4	976	17	BF300260	602032168
	199	17	9.4	978	14	BI818776	603037621
	200	17	9.4	982	17	BQ936776	AGENCOURT
c	201	17	9.4	990	12	BF233254	603062042
c	202	17	9.4	992	13	BI519758	Tetraodon
	203	17	9.4	1033	17	AL276794	Tetraodon
	204	17	9.4	1035	9	AL515482	AL515482
c	205	17	9.4	1077	14	BQ647831	AGENCOURT
	206	17	9.4	1153	17	AL401204	T3_end_of
	207	17	9.4	1176	14	BQ964998	AGENCOURT
	208	17	9.4	1252	17	BH537701	T7_end_of
c	209	17	9.4	2876	11	AL187522	Tetraodon
	210	17	9.4	3277	17	BG246761	602360839
c	211	17	9.4			AL410137	T7_end_of
c	212	17	9.4			AL413906	T3_end_of
	213	17	9.4			BQ935301	AGENCOURT
c	214	17	9.4			AY109642	Zea_mays
	215	17	9.4			BC026814	Mus_muscu
	216	16	8.8	108	9	AF219030	AF219030
	217	16	8.8	108	10	AJ398212	AJ398212
c	218	16	8.8	114	13	AW652251	100108_MA
	219	16	8.8	118	10	BI898910	480710_MA
	220	16	8.8	121	17	AW336599	22643_MAR
	221	16	8.8	151	9	AW336608	22631_MAR
c	222	16	8.8	156	13	AA691655	vs13b11.r
c	223	16	8.8	165	13	BM030805	495326_MA
c	224	16	8.8	166	9	BE600083	PI1_79_B0
c	225	16	8.8	181	10	AI925100	wn50a01_x
		16	8.8			BB151351	BB151351

Left column (rows 226–298):

	#			pos		accession	accession	name
	226	8.8	16	202	10	AW708363	AW708363	c6b09ne.f
	227	8.8	16	202	12	BE935138	BE935138	PM1-MT010
c	228	8.8	16	218	9	AI671298	AI671298	wb15c02.x
	229	8.8	16	222	10	BE236567	BE236567	145257 MA
	230	8.8	16	226	14	BQ848149	BQ848149	QGAGG15.y
	231	8.8	16	232	10	AV357477	AV357477	AV357477
c	232	8.8	16	234	14	T04951	T04951	EST02839 Fe
	233	8.8	16	237	12	BF043951	BF043951	BP250022B
	234	8.8	16	239	9	AV243375	AV243375	AV243375
	235	8.8	16	239	10	AW327165	AW327165	20806 MAR
	236	8.8	16	239	14	BQ739028	BQ739028	pt39e02.y
c	237	8.8	16	240	12	BF170784	BF170784	PCL1486 M
	238	8.8	16	241	14	T38552	T38552	EST104055 S
c	239	8.8	16	243	9	AA367182	AA367182	EST78228
	240	8.8	16	249	13	BJ270942	BJ270942	BJ270942
	241	8.8	16	252	12	BF602546	BF602546	267804 MA
	242	8.8	16	256	13	BM030794	BM030794	495310 MA
	243	8.8	16	256	14	N44172	N44172	yy29n08.r1
	244	8.8	16	263	12	BG889139	BG889139	EST514990
	245	8.8	16	264	17	A2929816	A2929816	479.di155
c	246	8.8	16	265	9	AV090954	AV090954	AV090954
c	247	8.8	16	266	10	AV358546	AV358546	AV358546
	248	8.8	16	270	10	BB595946	BB595946	BB595946
c	249	8.8	16	271	9	AA116915	AA116915	mn18g08.r
	250	8.8	16	271	14	AA346664	AA346664	EST52857
c	251	8.8	16	273	14	BQ296501	BQ296501	san93b02.
	252	8.8	16	280	10	AV298209	AV298209	AV298209
	253	8.8	16	283	13	BB554342	BB554342	BB554342
	254	8.8	16	283	12	BI556714	BI556714	603239435
	255	8.8	16	286	12	BF414001	BF414001	UI-R-BJ2-
	256	8.8	16	287	17	BH827709	BH827709	BACPP26-G
c	257	8.8	16	288	9	AV082848	AV082848	AV082848
	258	8.8	16	289	10	BB721138	BB721138	BB721138
	259	8.8	16	293	10	BB551744	BB551744	hx99b02.x
	260	8.8	16	294	10	BE666619	BE666619	150195 MA
	261	8.8	16	295	9	AV051555	AV051555	AV051555
c	262	8.8	16	295	12	BE752789	BE752789	205371 MA
c	263	8.8	16	300	9	AJ463752	AJ463752	AJ463752
	264	8.8	16	300	13	AU113054	AU113054	AU113054
c	265	8.8	16	308	13	BI050901	BI050901	CM4-GN016
	266	8.8	16	308	14	BI538045	BI538045	428408 MA
	267	8.8	16	313	9	BU004529	BU004529	QGGSH11.y
	268	8.8	16	315	17	AQ261842	AQ261842	CITBI-E1-
	269	8.8	16	315	10	AA347331	AA347331	EST53612
	270	8.8	16	317	10	AV307404	AV307404	AV307404
	271	8.8	16	319	17	BB208325	BB208325	BB208325
	272	8.8	16	319	9	AA911548	AA911548	oe78g06.s
	273	8.8	16	320	9	AA576259	AA576259	nm62c07.s
c	274	8.8	16	320	13	AI350592	AI350592	qq37d03.x
c	275	8.8	16	327	9	BM191144	BM191144	daj33b01.
	276	8.8	16	329	9	AA901384	AA901384	UI-R-A1-d
	277	8.8	16	329	10	AW432136	AW432136	73981 MAR
	278	8.8	16	337	10	BE404227	BE404227	WHE1203_F
	279	8.8	16	338	12	BF811848	BF811848	CM1-CIO13
c	280	8.8	16	338	13	BM161472	BM161472	EST563995
	281	8.8	16	341	17	BH822585	BH822585	BACPP17-F
	282	8.8	16	342	12	BF558660	BF558660	UI-R-A1-d
	283	8.8	16	345	12	AZ134610	AZ134610	OSJNBb011
	284	8.8	16	346	10	BB210412	BB210412	BB210412
	285	8.8	16	348	9	AI233060	AI233060	EST229748
	286	8.8	16	348	12	AU277265	AU277265	AU277265
	287	8.8	16	348	17	B61785	B61785	T19B11TF TA
	288	8.8	16	353	12	BF077533	BF077533	227463 MA
	289	8.8	16	356	9	AW682197	AW682197	EST009969
	290	8.8	16	358	9	AI100922	AI100922	EST210211
	291	8.8	16	358	13	AI819632	AI819632	wj87g01.x
	292	8.8	16	361	14	T77747	T77747	yd1d10.s1
	293	8.8	16	363	10	BG730736	BG730736	dac22c02.
c	294	8.8	16	367	12	BI285999	BI285999	UI-R-CX08
	295	8.8	16	369	13	AA912692	AA912692	o130d04.s
	296	8.8	16	374	17	AQ067789	AQ067789	HS_2239_B
c	297	8.8	16	375	14	W54813	W54813	md06h05.r1
	298	8.8	16	376	17	AQ296160	AQ296160	HS_2227_A

Right column (rows 299–371):

	#			pos		accession	accession	name
c	299	8.8	16	377	13	BI400566	BI400566	MI-P-AY1-
	300	8.8	16	379	16	AA869125	AA869125	vq4a11.r
c	301	8.8	16	379	10	AW707097	AW707097	sk21d11.y
	302	8.8	16	379	17	AQ238281	AQ238281	RPCI11-69
	303	8.8	16	380	17	AA509724	AA509724	vg20f04.r
c	304	8.8	16	383	17	AQ876759	AQ876759	HS_2114_B
	305	8.8	16	385	14	R00784	R00784	ye78f10.r1
	306	8.8	16	386	17	AQ794853	AQ794853	nbxb00541
	307	8.8	16	388	12	BE717358	BE717358	RC2-HT077
	308	8.8	16	389	17	AQ385514	AQ385514	RPCI11-13
	309	8.8	16	390	17	AU234292	AU234292	AU234292
	310	8.8	16	390	14	H77578	H77578	yu22f10.s1
	311	8.8	16	392	9	AI228997	AI228997	EST225692
	312	8.8	16	398	9	AI693574	AI693574	wd12c04.x
	313	8.8	16	398	10	AW550167	AW550167	L0G0H12-
	314	8.8	16	398	13	BI513075	BI513075	BB16001lA
	315	8.8	16	400	9	AU278695	AU278695	AU278695
	316	8.8	16	401	9	AA758083	AA758083	ah68a03.s
	317	8.8	16	401	9	AI018056	AI018056	ov64d11.s
	318	8.8	16	402	9	AA646638	AA646638	ve42c06.r
	319	8.8	16	402	10	AW611813	AW611813	hg81d01.x
	320	8.8	16	405	9	AI520767	AI520767	ti07c10.x
c	321	8.8	16	406	9	AW592053	AW592053	hf38a10.x
	322	8.8	16	406	17	AQ150922	AQ150922	HS_J110_B
c	323	8.8	16	407	14	BQ865034	BQ865034	QGC28I0I.
	324	8.8	16	410	17	AQ793182	AQ793182	HS_4674_A
	325	8.8	16	412	17	AQ208848	AQ208848	HS_J229_A
	326	8.8	16	414	17	AQ111716	AQ111716	CIT-HSP-2
	327	8.8	16	415	10	AW548253	AW548253	L0013F08-
	328	8.8	16	417	12	BE751953	BE751953	204217 MA
	329	8.8	16	418	10	AV665297	AV665297	AV665297
	330	8.8	16	418	14	BM836519	BM836519	K-EST0112
	331	8.8	16	418	10	BH601088	BH601088	BOGVB85TR
	332	8.8	16	419	10	BE222128	BE222128	hu06g02.x
	333	8.8	16	420	9	AI353454	AI353454	zeho521.s
	334	8.8	16	421	9	AI367659	AI367659	qv75a01.x
	335	8.8	16	425	9	AI715359	AI715359	UI-R-Y0-a
	336	8.8	16	426	14	BQ243084	BQ243084	TaE15019C
	337	8.8	16	429	10	AW864360	AW864360	PM4-SN001
	338	8.8	16	430	10	BB700138	BB700138	BB700138
	339	8.8	16	431	17	AZ049309	AZ049309	GSSBru086
	340	8.8	16	432	12	BE843776	BE843776	RCO-TN007
	341	8.8	16	432	13	BM088349	BM088349	501797 MA
	342	8.8	16	433	10	BE108998	BE108998	UI-R-BS1-
	343	8.8	16	434	10	AW015188	AW015188	UI-H-BI0p
	344	8.8	16	435	17	AA978922	AA978922	LD33127.5
	345	8.8	16	437	10	BE005320	BE005320	CM1-BN011
	346	8.8	16	438	12	AJ281058	AJ281058	4A3A-ABA-
	347	8.8	16	439	13	BI437381	BI437381	gc59h02.y
	348	8.8	16	440	12	BE843777	BE843777	RCO-TN007
	349	8.8	16	440	12	AQ360944	AQ360944	HS_5036_B
	350	8.8	16	445	9	AI281948	AI281948	qm26d05.x
	351	8.8	16	445	17	AI819169	AI819169	wj32b11.x
	352	8.8	16	445	12	BF596860	BF596860	su92e07.y
	353	8.8	16	448	13	BG606596	BG606596	WHE2958_E
	354	8.8	16	449	12	BF758270	BF758270	RC4-CT047
	355	8.8	16	449	12	AQ805751	AQ805751	HS_3175_A
	356	8.8	16	450	12	BF000656	BF000656	7h34c07.x
	357	8.8	16	451	10	AA227804	AA227804	zr56e09.r
	358	8.8	16	451	17	AW653557	AW653557	102303 MA
	359	8.8	16	451	13	BH836482	BH836482	PST154-2
	360	8.8	16	452	12	BG379882	BG379882	UI-R-CS0-
c	361	8.8	16	458	17	BH682311	BH682311	BOMFX35TR
	362	8.8	16	459	13	BG952309	BG952309	CM4-CT062
	363	8.8	16	459	13	BJ490096	BJ490096	BJ490096
	364	8.8	16	459	17	AZ590637	AZ590637	1M0400P14
	365	8.8	16	463	12	BG139201	BG139201	EST479655
c	366	8.8	16	463	17	AQ706874	AQ706874	HS_5548_B
	367	8.8	16	466	13	AA059521	AA059521	mj6ld12.r
c	368	8.8	16	468	9	BM352200	BM352200	ig83g07.y
c	369	8.8	16	468	11	AI701547	AI701547	we36b11.x
c	370	8.8	16	470	11	AY068413	AY068413	Schmidtea
	371	8.8	16	470	12	BF821102	BF821102	MR1-RT004

flag	#	len	score	#	cnt	accession	name
c	445	16	8.8	536	17	AQ671045	AQ671045 HS_5455_A
	446	16	8.8	537	14	BM900553	BM900553 rc38h09.y
c	447	16	8.8	541	9	AAS29404	AA529404 vi39h12.r
	448	16	8.8	543	13	BJ458369	BJ458369 BJ458369
c	449	16	8.8	545	12	BG265346	BG265346 1000021G0
	450	16	8.8	547	10	BB770016	BB770016 BB770016
	451	16	8.8	547	14	BM967114	BM967114 ij30f04.y
c	452	16	8.8	548	9	AA399714	AA399714 vd71g04.r
c	453	16	8.8	549	10	BE057349	BE057349 sn02a01.y
	454	16	8.8	549	10	BE237770	BE237770 894038B03
c	455	16	8.8	549	17	BH480224	BH480224 BOGOD65TR
	456	16	8.8	550	13	BM611383	BM611383 170006871
	457	16	8.8	551	10	AW677554	AW677554 110966_MA
	458	16	8.8	552	17	AQ267280	AQ267280 RPCI11-73
	459	16	8.8	552	17	AQ583875	AQ583875 RPCI-11-4
	460	16	8.8	555	10	AW736780	AW736780 JAYL0265.
c	461	16	8.8	555	13	BI846805	BI846805 fp88f07.y
	462	16	8.8	555	10	BQ302191	BQ302191 RCO-BT033
	463	16	8.8	555	14	TA115F03Q	AL465806 T._brucei
	464	16	8.8	556	17	BM142272	BM142272 if33b10.y
	465	16	8.8	557	13	BF482692	BF482692 WHE2301-2
	466	16	8.8	557	13	BI239025	BI239025 RE35409.5
	467	16	8.8	557	13	BI984990	BI984990 fu13c04.y
	468	16	8.8	557	17	AQ728661	AQ728661 HS_5479_A
	469	16	8.8	557	17	AZ258145	AZ258145 RPCI-23-1
	470	16	8.8	557	17	BH540246	BH540246 BOGKY08TF
	471	16	8.8	558	17	AQ724290	AQ724290 HS_2104_A
	472	16	8.8	560	9	AL729124	AL729124 AL729124
	473	16	8.8	561	13	BI228137	BI228137 RE25582.5
	474	16	8.8	562	10	AV607179	AV607179 AV607179
	475	16	8.8	562	10	AW632049	AW632049 91568_MAR
	476	16	8.8	562	14	BQ657375	BQ657375 HA08J23u
c	477	16	8.8	563	14	BM971138	BM971138 UI-CF-EC1
	478	16	8.8	564	10	AW871777	AW871777 da93h06.y
	479	16	8.8	564	17	BJ512071	BJ512071 BJ512071
	480	16	8.8	565	9	AI783103	AI783103 614008E10
	481	16	8.8	565	13	BJ515685	BJ515685 BJ515685
	482	16	8.8	566	10	AW509456	AW509456 si37d03.y
c	483	16	8.8	566	17	AQ267971	AQ267971 RPCI11-73
	484	16	8.8	567	13	BG988209	BG988209 MR1-HT106
c	485	16	8.8	567	17	AQ940976	AQ940976 Sheared_D
	486	16	8.8	568	10	BE444901	BE444901 WHE1129_F
c	487	16	8.8	568	14	BQ461508	BQ461508 HD04B21r
c	488	16	8.8	570	13	BJ510864	BJ510864 BJ510864
	489	16	8.8	570	17	BM291208	BM291208 CH230-103
	490	16	8.8	572	9	AI398233	AI398233 NCC3A6T7
c	491	16	8.8	573	10	BE679600	BE679600 df50b06.y
c	492	16	8.8	573	12	BF019215	BF019215 uy02e04.y
c	493	16	8.8	574	12	BF258117	BF258117 HVSMEf001
c	494	16	8.8	574	17	AQ312736	AQ312736 RPCI11-1K
	495	16	8.8	575	12	BF053866	BF053866 EST439012
	496	16	8.8	577	9	AI177005	AI177005 EST220612
c	497	16	8.8	577	10	BE033601	BE033601 MF04F10_M
c	498	16	8.8	577	17	AQ660287	AQ660287 Sheared_D
	499	16	8.8	578	9	AV150372	AV150372 AV150372
c	500	16	8.8	578	10	AW920039	AW920039 EST351343
c	501	16	8.8	579	13	BE443153	BE443153 WHE1113_A
	502	16	8.8	580	13	BI238343	BI238343 RE34768.5
	503	16	8.8	580	14	BQ659814	BQ659814 HF01M20w
	504	16	8.8	582	12	BF125940	BF125940 601762959
	505	16	8.8	582	13	BJ462563	BJ462563 BJ462563
c	506	16	8.8	583	13	BI239698	BI239698 RE36267.5
	507	16	8.8	583	17	BH338952	BH338952 CH230-89E
	508	16	8.8	584	9	AL677773	AL677773 AL677773
	509	16	8.8	584	10	BE455444	BE455444 HVSMEg001
c	510	16	8.8	585	12	AV932677	AV932677 AV932677
c	511	16	8.8	585	12	BF297688	BF297688 054PbH06
c	512	16	8.8	588	13	BJ500627	BJ500627 BJ500627
	513	16	8.8	588	12	BE922088	BE922088 EST425857
	514	16	8.8	589	17	AQ118850	AQ118850 HS_3016_B
	515	16	8.8	589	9	AU037872	AU037872 AU037872
	516	16	8.8	589	17	BH207239	BH207239 Sm1-55O16
c	517	16	8.8	590	12	BF285158	BF285158 EST449749

flag	#	len	score	#	cnt	accession	name
c	372	16	8.8	471	10	BB862657	BB862657 BB862657
c	373	16	8.8	472	9	AA868300	AA868300 ak40e01.s
	374	16	8.8	474	14	BJ072156	BJ072156 BJ072156
	375	16	8.8	475	12	N92747	N92747 zb26a03.s1
	376	16	8.8	477	13	BF595926	BF595926 su67f06.y
	377	16	8.8	479	9	AA285993	AA285993 vb88h01.r
	378	16	8.8	480	13	BJ507848	BJ507848 BJ507848
	379	16	8.8	480	9	AJ467389	AJ467389 AJ467389
c	380	16	8.8	480	9	AJ468389	AJ468389 AJ468389
c	381	16	8.8	480	9	AJ480753	AJ480753 AJ480753
c	382	16	8.8	480	9	AJ480754	AJ480754 AJ480754
	383	16	8.8	481	17	AQ207582	AQ207582 HS_2243_B
c	384	16	8.8	482	13	BJ243559	BJ243559 BJ243559
	385	16	8.8	484	17	AQ810982	AQ810982 HS_5379_A
	386	16	8.8	484	13	BI892529	BI892529 ea60g08.
c	387	16	8.8	485	9	BM372226	BM372226 EBro03_SQ
	388	16	8.8	485	9	AA938197	AA938197 oc10f01.s
c	389	16	8.8	485	9	AI104330	AI104330 EST213619
	390	16	8.8	486	10	AW017710	AW017710 614062C03
	391	16	8.8	487	14	R34734	R34734 yg61c08.r1
	392	16	8.8	491	9	BB701363	BB701363 BB701363
	393	16	8.8	491	14	AA309083	AA309083 EST180193
	394	16	8.8	491	14	C04902	C04902 C04902 Huma
	395	16	8.8	493	14	BM879973	BM879973 ku03d05.y
	396	16	8.8	493	14	BQ446019	BQ446019 UI-H-EU1-
	397	16	8.8	493	17	AQ683622	AQ683622 HS_5458_B
	398	16	8.8	494	17	BH700692	BH700692 BOMMG89TR
	399	16	8.8	499	9	AI701172	AI701172 we10h03.x
c	400	16	8.8	502	17	AQ345230	AQ345230 RPCI11-13
c	401	16	8.8	503	13	BE234164	BE234164 140909_MA
c	402	16	8.8	503	13	BJ477589	BJ477589 BJ477589
c	403	16	8.8	504	10	BE628622	BE628622 uu03d03.x
c	404	16	8.8	504	17	BH858605	BH858605 S3_008a_t
c	405	16	8.8	506	13	AQ568202	AQ568202 HS_5233_B
c	406	16	8.8	506	13	BJ505842	BJ505842 BJ505842
	407	16	8.8	506	13	BM142258	BM142258 if33a03.y
	408	16	8.8	506	14	BM680634	BM680634 UI-E-EJO-
c	409	16	8.8	507	14	BM728073	BM728073 UI-E-EJO-
c	410	16	8.8	508	12	N47753	N47753 yy55d03.r1
c	411	16	8.8	510	13	BG224862	BG224862 kp51e02.y
	412	16	8.8	510	10	FR0004797	Z88607 F.rubripes
	413	16	8.8	511	10	BE135082	BE135082 ug22f09.y
	414	16	8.8	512	17	AQ557622	AQ557622 HS_2083_A
	415	16	8.8	512	14	BM819930	BM819930 K-EST0088
c	416	16	8.8	513	17	AQ366297	AQ366297 nbxb00065H
	417	16	8.8	513	17	AZ019690	AZ019690 RPCI-23-2
c	418	16	8.8	516	17	AQ412035	AQ412035 RPCI-11-1
c	419	16	8.8	517	17	AZ929245	AZ929245 479.dif23
c	420	16	8.8	518	9	AI603341	AI603341 UI-R-AE0-
	421	16	8.8	519	9	AI456841	AI456841 LD36988.5
	422	16	8.8	519	17	AQ709145	AQ709145 HS_5329_B
	423	16	8.8	519	14	AQ896765	AQ896765 HS_5526_B
c	424	16	8.8	520	10	BE077132	BE077132 RC5-BT060
c	425	16	8.8	521	14	BQ557447	BQ557447 H4046B10-
	426	16	8.8	522	13	BI374080	BI374080 RE61593.5
c	427	16	8.8	523	9	AA679306	AA679306 zi28d01.s
c	428	16	8.8	523	10	AW065819	AW065819 614062C03
c	429	16	8.8	523	17	AQ595545	AQ595545 HS_5453_A
	430	16	8.8	526	12	BG764292	BG764292 602736169
	431	16	8.8	527	9	AI927079	AI927079 wo87g04.x
c	432	16	8.8	527	17	AQ690797	AQ690797 nbxb00083A
c	433	16	8.8	528	17	AQ433737	AQ433737 HS_5063_A
c	434	16	8.8	529	17	BG692457	BG692457 342152_BA
	435	16	8.8	529	14	BM900564	BM900564 rc39a08.y
	436	16	8.8	531	10	AQ799898	AQ799898 HS_5291_A
c	437	16	8.8	531	13	BE077129	BE077129 RC5-BT060
	438	16	8.8	531	13	BJ500627	BJ500627 BJ500627
c	439	16	8.8	532	13	BI977137	BI977137 486436_MA
	440	16	8.8	532	13	BJ514276	BJ514276 BJ514276
c	441	16	8.8	533	17	AQ449124	AQ449124 mgxb0022P
c	442	16	8.8	534	14	BM900845	BM900845 rc42b10.y
c	443	16	8.8	534	17	TA36A10Q	AL494921 T._brucei
	444	16	8.8	535	9	AI783104	AI783104 614008E11

Left block:

	Index			Pos	N	Accession	Name
	518	8.8	16	590	13	BI876631	BI876631 f174f10.y
	519	8.8	16	590	17	AQ808093	AQ808093 HS_2169_B
c	520	8.8	16	591	9	AL450836	AL450836 AL450836
	521	8.8	16	591	13	BJ496570	BJ496570 BJ496570
c	522	8.8	16	591	14	BQ464572	BQ464572 HF02K15r
	523	8.8	16	592	17	FR0032133	FR0032133 Fugu rubr
	524	8.8	16	593	10	AW000034	AW000034 614008B10
	525	8.8	16	594	13	BI706704	BI706704 fq09a12.y
	526	8.8	16	594	17	AQ614138	AQ614138 HS_5127_B
	527	8.8	16	595	13	BF035077	BF035077 601456180
	528	8.8	16	595	17	BJ463301	BJ463301 BJ463301
c	529	8.8	16	595	13	AQ452610	AQ452610 HS_5184_A
c	530	8.8	16	596	17	AQ788766	AQ788766 HS_3187_B
	531	8.8	16	597	13	BJ493764	BJ493764 BJ493764
c	532	8.8	16	598	12	BG087706	BG087706 H3142C08-
	533	8.8	16	600	12	BF297871	BF297871 057PbC10
c	534	8.8	16	602	12	BG474589	BG474589 602517378
	535	8.8	16	603	13	BJ491153	BJ491153 BJ491153
	536	8.8	16	603	13	BJ508296	BJ508296 BJ508296
	537	8.8	16	604	9	AU268608	AU268608 AU268608
	538	8.8	16	605	14	BQ852936	BQ852936 QGB19G06.
	539	8.8	16	608	10	BE305749	BE305749 601102483
	540	8.8	16	608	14	BJ859473	BJ859473 QGC13C04.
c	541	8.8	16	610	9	AA869941	AA869941 vq19f04.r
	542	8.8	16	610	10	BE237771	BE237771 894038B03
	543	8.8	16	611	13	BM013735	BM013735 603639282
c	544	8.8	16	611	17	BH564652	BH564652 BOGTN69TF
	545	8.8	16	614	13	BI670857	BI670857 fp57d01.x
	547	8.8	16	614	14	BU008823	BU008823 QGH8K21.Y
	547	8.8	16	615	12	BF823531	BF823531 MR1-RT003
	548	8.8	16	616	9	AU268230	AU268230 AU268230
c	549	8.8	16	617	14	BQ568037	BQ568037 gi100f10.
c	550	8.8	16	617	14	BQ874297	BQ874297 QGI4O24.y
	551	8.8	16	618	10	AW961697	AW961697 EST373770
	552	8.8	16	619	12	BG414709	BG414709 HVSMEk000
	553	8.8	16	620	13	BJ516894	BJ516894 BJ516894
c	554	8.8	16	620	14	BM969585	BM969585 UI-CF-DU1
	555	8.8	16	621	17	B89851	B89851 CIT-HSP-216
c	556	8.8	16	621	12	BF019225	BF019225 uy02f04.y
	557	8.8	16	621	14	BQ579072	BQ579072 WHE2965_B
	558	8.8	16	622	10	AW561920	AW561920 IPPGH2O00
	559	8.8	16	623	17	BH488309	BH488309 BOGTG10TF
	560	8.8	16	624	12	BG337484	BG337484 602435016
	561	8.8	16	627	14	BF219769	BF219769 601296696
c	562	8.8	16	627	14	BQ763425	BQ763425 EBro02_SQ
	563	8.8	16	627	14	BQ873997	BQ873997 QGI4C03.y
c	564	8.8	16	628	10	AV706254	AV706254 AV706254
	565	8.8	16	628	13	BI243303	BI243303 RE40888.5
c	566	8.8	16	628	17	AZ345044	AZ345044 1M0079H23
	567	8.8	16	630	13	BH524153	BH524153 BOHBYO7TR
	568	8.8	16	631	12	BG769478	BG769478 602742443
	569	8.8	16	633	10	BE299981	BE299981 600944553
	570	8.8	16	633	13	AQ978714	AQ978714 RPCI-23-2
	571	8.8	16	634	17	AI456135	AI456135 LD35990.5
c	572	8.8	16	634	10	BE253472	BE253472 601108146
	573	8.8	16	634	12	BG307382	BG307382 fm03e01.y
c	574	8.8	16	634	12	BF298070	BF298070 05PbH03
	575	8.8	16	634	17	BJ312150	BJ312150 BJ312150
c	576	8.8	16	636	13	BJ493536	BJ493536 BJ493536
	577	8.8	16	639	13	BJ499765	BJ499765 BJ499765
	578	8.8	16	639	17	B61853	B61853 T19G10TF TA
c	579	8.8	16	640	13	BI401433	BI401433 MI-P-CP0-
	580	8.8	16	641	13	BJ517160	BJ517160 BJ517160
	581	8.8	16	642	12	BF984087	BF984087 602307552
	582	8.8	16	643	17	BJ513560	BJ513560 BJ513560
	583	8.8	16	647	9	AI533183	AI533183 SD04924.5
	584	8.8	16	647	13	BJ045190	BJ045190 BJ045190
	585	8.8	16	647	13	BJ275892	BJ275892 BJ275892
	586	8.8	16	647	13	BJ517590	BJ517590 BJ517590
c	587	8.8	16	648	9	AI325107	AI325107 mr80a04.x
	588	8.8	16	649	13	AA140803	AA140803 CK00476.5
	589	8.8	16	649	13	BJ239606	BJ239606 BJ239606
c	590	8.8	16	651	10	AV700784	AV700784 AV700784

Right block:

	Index			Pos	N	Accession	Name
c	591	16	8.8	652	9	AI171951	AI171951 EST217945
c	592	16	8.8	652	12	BF296903	BF296903 044PbC07
	593	16	8.8	652	13	BJ513773	BJ513773 BJ513773
	594	16	8.8	653	14	BM890463	BM890463 fy6h08.y
	595	16	8.8	653	17	AZ339768	AZ339768 1M0071E20
c	596	16	8.8	654	17	AZ937809	AZ937809 2M0196E12
	597	16	8.8	654	12	BF967585	BF967585 602287443
c	598	16	8.8	654	13	BE840964	BE840964 PM4-SN001
c	599	16	8.8	654	13	BJ493315	BJ493315 BJ493315
c	600	16	8.8	654	13	BJ495850	BJ495850 BJ495850
	601	16	8.8	655	13	BJ498716	BJ498716 BJ498716
c	602	16	8.8	655	13	BJ270887	BJ270887 BJ270887
c	603	16	8.8	656	14	BQ459599	BQ459599 HA08J23r
	604	16	8.8	657	13	BJ488314	BJ488314 BJ488314
	605	16	8.8	659	13	AZ276418	AZ276418 RPCI-23-1
	606	16	8.8	659	13	BJ511554	BJ511554 BJ511554
	607	16	8.8	659	14	BQ907233	BQ907233 N005F10 O
c	608	16	8.8	661	13	BH402516	BH402516 AG-ND-172
c	609	16	8.8	661	13	BJ498835	BJ498835 BJ498835
	610	16	8.8	661	13	BJ515798	BJ515798 BJ515798
	611	16	8.8	662	13	BJ273368	BJ273368 BJ273368
	612	16	8.8	662	13	BJ493890	BJ493890 BJ493890
	613	16	8.8	664	17	BH707006	BH707006 BOMCN47TR
	614	16	8.8	665	13	BJ515018	BJ515018 BJ515018
	615	16	8.8	665	14	BQ864314	BQ864314 QGC26F24.
c	616	16	8.8	666	17	AG021640	AG021640 Oryza sat
	617	16	8.8	669	17	BQ410415	BQ410415 GA_Ed003
c	618	16	8.8	670	12	BE915603	BE915603 601668650
	619	16	8.8	672	17	AQ926930	AQ926930 RPCI-23-3
	620	16	8.8	674	13	BJ503554	BJ503554 BJ503554
	621	16	8.8	674	17	BH723370	BH723370 BOMCFO1TR
c	622	16	8.8	676	17	AG179916	AG179916 Pan trogl
c	623	16	8.8	677	13	BQ100239	BQ100239 ph74c08.y
	624	16	8.8	678	14	BQ284204	BQ284204 f28a10.y
	625	16	8.8	678	17	BQ506924	BQ506924 EST614339
	626	16	8.8	679	9	AU138797	AU138797 AU138797
	627	16	8.8	680	12	BF505653	BF505653 AT07924.5
	628	16	8.8	682	14	BH526223	BH526223 BOGZC90TF
	629	16	8.8	683	14	BQ874562	BQ874562 QGI5J20.y
	630	16	8.8	684	13	BJ511011	BJ511011 BJ511011
	631	16	8.8	685	17	AZ358026	AZ358026 1M0100C01
	632	16	8.8	686	13	BJ503905	BJ503905 BJ503905
	633	16	8.8	686	17	BJ517459	BJ517459 BJ517459
	634	16	8.8	686	17	AG096607	AG096607 Pan trogl
	635	16	8.8	687	13	BJ051834	BJ051834 BJ051834
	636	16	8.8	687	13	BJ518288	BJ518288 BJ518288
	637	16	8.8	689	13	BQ616715	BQ616715 faa94c06.
	638	16	8.8	689	14	AQ381905	AQ381905 RPCI11-13
	639	16	8.8	690	12	BF575620	BF575620 602133018
	640	16	8.8	692	13	BJ517767	BJ517767 BJ517767
	641	16	8.8	693	13	BJ505612	BJ505612 BJ505612
	642	16	8.8	693	17	BH517222	BH517222 BOGUIO2TR
	643	16	8.8	694	17	AQ717743	AQ717743 HS_5507_B
	644	16	8.8	695	10	BB396051	BB396051 601312669
	645	16	8.8	695	13	BJ509831	BJ509831 BJ509831
	646	16	8.8	696	14	BM713866	BM713866 UI-E-EJ0-
c	647	16	8.8	696	14	BQ851370	BQ851370 QGB15C11.
	648	16	8.8	700	12	BG417262	BG417262 HVSMEK001
c	649	16	8.8	700	12	BG858159	BG858159 1024056B0
	650	16	8.8	700	13	BJ487559	BJ487559 BJ487559
c	651	16	8.8	701	13	BJ516415	BJ516415 BJ516415
	652	16	8.8	701	13	BJ505730	BJ505730 BJ505730
	653	16	8.8	701	13	BJ507555	BJ507555 BJ507555
	654	16	8.8	702	13	BJ511553	BJ511553 BJ511553
	655	16	8.8	702	13	BJ515887	BJ515887 BJ515887
c	656	16	8.8	703	9	AL729189	AL729189 AL729189
	657	16	8.8	703	13	AV732669	AV732669 AV732669
c	658	16	8.8	703	10	BB023294	BB023294 BB023294
	659	16	8.8	703	10	BM657125	BM657125 170006874
c	660	16	8.8	704	10	AW864458	AW864458 PM4-SN001
	661	16	8.8	704	13	BJ508682	BJ508682 BJ508682
c	662	16	8.8	705	14	BQ540236	BQ540236 PTAM0553
c	663	16	8.8	706	13	BJ514245	BJ514245 BJ514245

Left block:

c	idx			pos	len	acc	name
c	664	8.8	16	706	13	BJ518418	BJ518418
c	665	8.8	16	706	14	BQ473046	rc82h08.y
	666	8.8	16	707	9	AL714645	AL714645
	667	8.8	16	707	10	BG631938	AT26734.5
	668	8.8	16	709	10	AV835697	AV835697
	669	8.8	16	709	12	BF502369	AT18070.5
	670	8.8	16	710	14	BQ863972	QGC25G03.
	671	8.8	16	711	17	BH575550	BOHGF16TR
	672	8.8	16	713	13	BJ502744	BJ502744
	673	8.8	16	714	17	BE266236	BE266236
	674	8.8	16	715	9	AU142709	AU142709
c	675	8.8	16	715	13	BJ303604	BJ303604
c	676	8.8	16	715	17	BH029477	RPCI-24-2
	677	8.8	16	716	9	AL516179	AL516179
	678	8.8	16	716	10	BE602727	HVSMEh010
c	679	8.8	16	717	12	BQ507720	EST615135
	680	8.8	16	717	9	AI950968	wx6if07.x
	681	8.8	16	717	13	BJ516453	BJ516453
	682	8.8	16	720	10	BE601716	HVSMEh009
	683	8.8	16	721	10	AV761290	AV761290
	684	8.8	16	722	13	BE602268	HVSMEh009
c	685	8.8	16	724	12	BI545877	603188065
	686	8.8	16	724	17	BH739269	BOMNU49TR
	687	8.8	16	724	17	AG169650	Pan trogl
c	688	8.8	16	727	10	AW421973	fi89c03.-y
c	689	8.8	16	729	9	AL714621	AL714621
	690	8.8	16	729	13	BI829891	603080003
c	691	8.8	16	730	13	BH597370	BOGO187TR
	692	8.8	16	731	13	BI834542	603089731
	693	8.8	16	731	17	AG175563	Pan trogl
	694	8.8	16	733	17	AG097715	Pan trogl
	695	8.8	16	735	10	BE602520	HVSMEh009
	696	8.8	16	735	12	BE791393	BE791393
	697	8.8	16	738	12	BF298823	021PDG0B
	698	8.8	16	738	17	BH601957	BOGKO03TF
	699	8.8	16	740	17	AQ423741	CITB1-E1-
	700	8.8	16	741	9	AA391681	LD11614.5
	701	8.8	16	743	17	AF010666	AF010666
	702	8.8	16	745	10	BI772769	603059369
	703	8.8	16	747	12	AV757155	AV757155
	704	8.8	16	748	12	BG763482	602735569
	705	8.8	16	748	17	BH661738	BOH3X71TR
	706	8.8	16	748	17	AG126496	Pan trogl
	707	8.8	16	749	17	AG043497	Pan trogl
	708	8.8	16	750	12	BJ306972	BJ306972
	709	8.8	16	750	14	BQ804524	BQ804524
	710	8.8	16	752	17	AQ935716	CpG2613A
	711	8.8	16	757	13	BG592245	EST500087
	712	8.8	16	758	17	BI259674	602968585
	713	8.8	16	758	17	AQ865327	nbeb0025K
	714	8.8	16	760	17	BH672363	BOMMT65TF
c	715	8.8	16	760	17	AG185520	Pan trogl
	716	8.8	16	765	12	BF489011	AT24695.5
	717	8.8	16	767	17	BQ483259	WHE3506.D
	718	8.8	16	769	12	BG176742	602313562
	719	8.8	16	769	13	BJ504088	BJ504088
	720	8.8	16	770	17	BH589606	BOGIL42TF
c	721	8.8	16	770	17	AZ186957	SP_1008_A
	722	8.8	16	771	9	AU138884	AU138884
	723	8.8	16	771	17	BH501851	BOGKW15TR
c	724	8.8	16	772	17	BJ428639	BJ428639
	725	8.8	16	775	17	BH439653	BOGBI59TF
	726	8.8	16	776	12	BG682402	602630142
	727	8.8	16	778	17	BH246876	BOGAP46TR
	728	8.8	16	778	17	BH477252	BOGSR53TF
	729	8.8	16	779	17	AZ391357	1M0153H19
	730	8.8	16	780	12	BF723848	EST00661
c	731	8.8	16	780	12	BF723929	EST00827
	732	8.8	16	780	12	BG189705	RST8751_A
	733	8.8	16	781	10	BE229825	98SN0215
	734	8.8	16	783	12	BG529115	602579044
	735	8.8	16	784	17	AU133065	AU133065
c	736	8.8	16	785	10	BE265716	601193729

Right block:

c	idx			pos	acc	name	acc2	len
c	737	8.8	16	785	BG199942	RST19238	BG199942	12
c	738	8.8	16	785	BG598450	EST503350	BG598450	12
c	739	8.8	16	785	BF162411	601771536	BF162411	12
c	740	8.8	16	788	BH651173	BOHTV27TF	BH651173	17
	741	8.8	16	789	AU138757	AU138757	AU138757	9
	742	8.8	16	789	BG338540	602436269	BG338540	12
	743	8.8	16	789	AL235684	Tetraodon	CNS03AUJ	17
	744	8.8	16	790	AG172174	Pan trogl	AG172174	17
	745	8.8	16	791	BG211907	RST31482	BG211907	12
	746	8.8	16	791	BH096118	RPCI-24-2	BH096118	17
	747	8.8	16	792	BH662286	BOMOU58TR	BH662286	17
c	748	8.8	16	794	BE541664	601067757	BE541664	10
	749	8.8	16	794	BF672148	602150308	BF672148	12
	750	8.8	16	796	BH311828	CH230-26I	BH311828	17
c	751	8.8	16	800	BE281012	601159078	BE281012	10
	752	8.8	16	801	AZ713573	RPCI-24-1	AZ713573	17
	753	8.8	16	802	BH506565	BOGKQ19TR	BH506565	17
	754	8.8	16	803	BF670849	602149932	BF670849	12
	755	8.8	16	804	BH716051	BOMEW59TR	BH716051	17
	756	8.8	16	805	BG784988	SEAUMC004	BG784988	12
c	757	8.8	16	808	BG418377	HVSMEk002	BG418377	12
	758	8.8	16	808	BG780750	SEAUMC000	BG780750	12
	759	8.8	16	808	BG985781	5114_NICH	BG985781	13
c	760	8.8	16	809	BE745374	601578006	BE745374	12
	761	8.8	16	810	BG169153	602320622	BG169153	12
c	762	8.8	16	812	BG479932	BOGDO54TR	BG479932	17
	763	8.8	16	813	BG784920	SEAUMC004	BG784920	12
	764	8.8	16	815	BG286612	602381623	BG286612	12
	765	8.8	16	815	BH485748	BOGQY04TR	BH485748	17
	766	8.8	16	816	BH709010	BOHYR74TF	BH709010	17
	767	8.8	16	816	BH731281	BOMQF45TR	BH731281	17
c	768	8.8	16	818	BF131986	601820973	BF131986	12
	769	8.8	16	819	AL521903	AL521903	AL521903	9
c	770	8.8	16	822	BG438051	602490312	BG438051	12
c	771	8.8	16	822	BH499431	BOHEA73TR	BH499431	17
	772	8.8	16	831	BH652178	BOHSN47TF	BH652178	17
	773	8.8	16	834	BE537792	601067563	BE537792	10
	774	8.8	16	834	BF576426	602133838	BF576426	12
	775	8.8	16	840	BH061228	RPCI-24-3	BH061228	17
	776	8.8	16	842	BF246691	601185482	BF246691	12
	777	8.8	16	842	AL559482	AL559482	AL559482	9
	778	8.8	16	844	BI335393	602997963	BI335393	13
	779	8.8	16	844	BH669061	BOHUM34TR	BH669061	17
	780	8.8	16	848	AG127562	Pan trogl	AG127562	17
	781	8.8	16	852	BG176971	602313659	BG176971	12
	782	8.8	16	856	BF691156	602247244	BF691156	12
	783	8.8	16	857	BF574178	602131337	BF574178	12
	784	8.8	16	861	BH657646	BOMJN84TF	BH657646	17
	785	8.8	16	867	BG611573	602613487	BG611573	12
	786	8.8	16	872	BG577162	602599724	BG577162	12
	787	8.8	16	873	AL521387	AL521387	AL521387	9
	788	8.8	16	875	BI521362	603081147	BI521362	13
	789	8.8	16	877	BG774379	602662419	BG774379	12
c	790	8.8	16	877	AL200588	Tetraodon	CNS02JRN	17
c	791	8.8	16	877	BF672604	602152186	BF672604	12
	792	8.8	16	877	AZ550532	ENTG161TF	AZ550532	17
	793	8.8	16	878	AZ685522	ENTKS12TR	AZ685522	17
	794	8.8	16	879	AZ541085	ENTCZ30TR	AZ541085	17
	795	8.8	16	885	BH153447	ENTSA75TF	BH153447	17
	796	8.8	16	886	BF790868	602250665	BF790868	12
	797	8.8	16	890	BQ961372	602257502	BG479939	14
	798	8.8	16	890	BE792040	ENTDF08TR	AZ540086	12
	799	8.8	16	893	BG417662	601581729	BE792040	12
	800	8.8	16	895	BI765263	HVSMEk001	BG417662	13
	801	8.8	16	895	AZ679050	603050359	BI765263	17
	802	8.8	16	901	BQ278222	ENTM113TF	AZ679050	14
	803	8.8	16	901	BQ642905	AGENCOURT	BQ278222	14
	804	8.8	16	902		AGENCOURT		
	805	8.8	16	903		AL577389	AL577389	
	806	8.8	16	903		602734260	BG770528	
	807	8.8	16	903		AGENCOURT	BQ227266	
	808	8.8	16	904		AGENCOURT		
	809	8.8	16	904		AGENCOURT		

c	idx			pos	n	ID	ID / description
c	810	8.8	16	905	12	BF980093	BF980093 602288688
c	811	8.8	16	910	17	AZ676540	AZ676540 ENTJN10TF
c	812	8.8	16	911	12	BF972351	BF972351 602240664
c	813	8.8	16	913	12	BF137620	BF137620 601783146
c	814	8.8	16	913	17	AZ194200	AZ194200 SP_1025_B
c	815	8.8	16	916	17	CNS04RSW	CNS04RSW Tetraodon
c	816	8.8	16	917	12	CNS06DZI	CNS06DZI T3 end of
	817	8.8	16	919	12	BF688581	BF688581 602185007
	818	8.8	16	920	17	AZ693086	AZ693086 ENTHL72TR
c	819	8.8	16	920	17	AZ540221	AZ540221 ENTGJ45TR
c	820	8.8	16	920	17	AZ547298	AZ547298 ENTDM56TR
	821	8.8	16	922	17	BH151517	BH151517 ENTPW03TR
	822	8.8	16	924	14	BQ653455	BQ653455 AGENCOURT
	823	8.8	16	925	17	AL232264	AL232264 Tetraodon
	824	8.8	16	927	12	BF982882	BF982882 602305651
	825	8.8	16	934	17	BH139426	BH139426 ENTNW77TR
c	826	8.8	16	936	12	BG494503	BG494503 602539527
	827	8.8	16	938	12	BE883114	BE883114 601508432
	828	8.8	16	938	9	AU079202	AU079202 AU079202
	829	8.8	16	940	9	AL524921	AL524921 AL524921
	830	8.8	16	943	9	AL528487	AL528487 AL528487
	831	8.8	16	944	12	BF206280	BF206280 601869296
c	832	8.8	16	945	12	BG177668	BG177668 602314668
c	833	8.8	16	945	12	BG618168	BG618168 602645084
	834	8.8	16	945	17	AG128710	AG128710 Pan trogl
	835	8.8	16	946	14	BQ948317	BQ948317 AGENCOURT
	836	8.8	16	948	12	BG174610	BG174610 602334453
	837	8.8	16	949	12	BF972261	BF972261 602240743
c	838	8.8	16	949	17	CNS03IEL	AL245478 Tetraodon
	839	8.8	16	950	12	AL529674	AL529674 AL529674
	840	8.8	16	952	12	BE867343	BE867343 601442421
	841	8.8	16	953	17	CNS02NUY	AL205891 Tetraodon
c	842	8.8	16	955	9	AL528524	AL528524 AL528524
	843	8.8	16	956	14	BQ944343	BQ944343 AGENCOURT
c	844	8.8	16	958	9	AL531581	AL531581 AL531581
	845	8.8	16	958	14	BQ221659	BQ221659 AGENCOURT
	846	8.8	16	960	12	BF684463	BF684463 602142949
	847	8.8	16	972	9	AL561431	AL561431 AL561431
	848	8.8	16	973	9	AL525456	AL525456 AL525456
c	849	8.8	16	978	14	BQ920167	BQ920167 AGENCOURT
c	850	8.8	16	983	17	BH135005	BH135005 ENTNZ65TR
	851	8.8	16	984	14	BQ952335	BQ952335 AGENCOURT
c	852	8.8	16	993	17	AF010646	AF010646 AF010646
	853	8.8	16	996	9	AL530315	AL530315 AL530315
	854	8.8	16	996	17	CNS046XN	AL277268 Tetraodon
	855	8.8	16	1004	13	BM463246	BM463246 AGENCOURT
c	856	8.8	16	1010	17	CNS06P01	AL408743 T3 end of
	857	8.8	16	1011	13	BI948133	BI948133 HVSME1000
c	858	8.8	16	1012	12	BG109876	BG109876 602279544
	859	8.8	16	1017	12	BF798590	BF798590 601581621
	860	8.8	16	1030	13	BM542964	BM542964 AGENCOURT
	861	8.8	16	1030	12	BG391905	BG391905 602409835
	862	8.8	16	1034	13	BM468533	BM468533 AGENCOURT
	863	8.8	16	1043	17	BH737498	BH737498 BOMNI43TR
	864	8.8	16	1049	13	BM454306	BM454306 AGENCOURT
c	865	8.8	16	1051	17	CNS078GS	AL433970 T3 end of
c	866	8.8	16	1071	14	BM914900	BM914900 AGENCOURT
	867	8.8	16	1075	12	BF973320	BF973320 602242248
	868	8.8	16	1081	17	CNS054017	BM554017 Tetraodon
	869	8.8	16	1084	17	CNS02UT6	AL214899 Tetraodon
	870	8.8	16	1093	12	BG860087	BG860087 1024067D0
	871	8.8	16	1117	17	B08285	B08285 F7G20-Sp6.1
c	872	8.8	16	1156	13	BM465045	BM465045 AGENCOURT
c	873	8.8	16	1156	13	BM562235	BM562235 AGENCOURT
	874	8.8	16	1194	13	BG869368	BG869368 602788892
	875	8.8	16	1201	12	BF794404	BF794404 602255929
	876	8.8	16	1210	12	BF123292	BF123292 601759580
	877	8.8	16	1301	13	BF665567	BF665567 603287843
c	878	8.8	16	1353	12	BG846995	BG846995 1024015El
c	879	8.8	16	1461	14	BQ228523	BQ228523 AGENCOURT
c	880	8.8	16	1536	12	BQ208258	BQ208258 601871359
	881	8.8	16	1613	10	BE616058	BE616058 601279546
	882	8.8	16	1911	12	BF972146	BF972146 602240188

c	idx		pos	n	ID	ID / description
c	883	8.8	2841	17	BH770960	BH770960 LLMGtag68
c	884	8.8	2956	11	BC011390	BC011390 Mus muscu
c	885	8.3	76	9	AA976956	AA976956 oq26e05.s
	886	8.3	84	9	AA175868	AA175868 ms84g08.r
c	887	8.3	93	9	AA607260	AA607260 vm92907.r
	888	8.3	102	17	AZ814612	AZ814612 2MO082K20
	889	8.3	111	17	AL756809	AL756809 Arabidops
c	890	8.3	112	14	BQ095551	BQ095551 kk06f07.y
	891	8.3	119	17	AA235479	AA235479 zt35z02.s
	892	8.3	122	10	AW435403	AW435403 UI-R-BJ0p
	893	8.3	123	17	BH853382	BH853382 SALK_0769
c	894	8.3	123	17	BH853390	BH853390 SALK_0769
	895	8.3	125	13	BM146009	BM146009 TCAAP1D95
c	896	8.3	134	9	AV115659	AV115659 AV115659
	897	8.3	135	9	AV128518	AV128518 AV128518
c	898	8.3	135	9	AV131509	AV131509 AV131509
	899	8.3	135	10	AW846570	AW846570 QV0-CT017
	900	8.3	135	14	D18261	D18261 MUSGS00345
c	901	8.3	139	17	AQ061735	AQ061735 CIT-HSP-2
	902	8.3	141	17	AL764963	AL764963 Arabidops
	903	8.3	148	9	AI837034	AI837034 UI-M-AP0-
c	904	8.3	148	17	AA589299	AA589299 vm36a11.r
	905	8.3	148	13	BG956511	BG956511 PM4-CT080
c	906	8.3	149	12	BE832174	BE832174 PM3-MT007
c	907	8.3	151	12	BF069343	BF069343 st46c09.y
	908	8.3	156	9	AV130646	AV130646 AV130646
	909	8.3	164	17	A2895402	A2895402 RPCI-24-1
c	910	8.3	165	9	AV151065	AV151065 AV151065
	911	8.3	165	10	AW517685	AW517685 x86d04.x
	912	8.3	165	13	BM448431	BM448431 DSA011A07
	913	8.3	165	14	N93573	N93573 zb56g05.s1
c	914	8.3	166	9	AI014389	AI014389 ou93a10.s
	915	8.3	168	9	AV072628	AV072628 AV072628
c	916	8.3	170	9	AV123094	AV123094 AV123094
	917	8.3	170	9	AA245703	AA245703 mx30d07.r
	918	8.3	171	9	AA125042	AA125042 mq73f06.r
c	919	8.3	176	9	AV053042	AV053042 AV053042
	920	8.3	176	9	AV159284	AV159284 AV159284
	921	8.3	177	9	AV098193	AV098193 AV098193
c	922	8.3	180	12	BG626409	BG626409 cC-esflcl
	923	8.3	184	14	N84442	N84442 KK809F Hum
c	924	8.3	185	9	AV133346	AV133346 AV133346
	925	8.3	187	9	AV123166	AV123166 AV123166
	926	8.3	187	14	W46022	W46022 mc79a05.r1
c	927	8.3	188	9	AV057727	AV057727 AV057727
c	928	8.3	190	10	BB191043	BB191043 BB191043
	929	8.3	191	9	AV039816	AV039816 AV039816
	930	8.3	191	9	AV054749	AV054749 AV054749
	931	8.3	192	10	AW151324	AW151324 xe73a02.x
	932	8.3	192	13	BG956174	BG956174 QVi-CT075
c	933	8.3	194	9	AV132147	AV132147 AV132147
	934	8.3	195	9	AV030429	AV030429 AV030429
	935	8.3	197	9	AJ282322	AJ282322 4A3A-AAS-
	936	8.3	197	9	AV164024	AV164024 AV164024
c	937	8.3	198	9	AI712809	AI712809 UI-R-AF1-
c	938	8.3	199	9	AU012762	AU012762 AU012762
	939	8.3	199	9	AU012956	AU012956 AU012956
	940	8.3	199	9	AV057691	AV057691 AV057691
	941	8.3	199	10	BB000468	BB000468 BB000468
	942	8.3	200	9	AV088248	AV088248 AV088248
	943	8.3	201	9	AV049097	AV049097 AV049097
c	944	8.3	202	9	AI836756	AI836756 UI-M-AJ0-
	945	8.3	203	9	AJ207701	AJ207701 4A3A-AAS-
c	946	8.3	204	14	BQ974509	BQ974509 QHI15O16.
	947	8.3	207	9	AA740166	AA740166 ob26c06.s
	948	8.3	207	9	AV035503	AV035503 AV035503
	949	8.3	207	13	BI190429	BI190429 ildl2fs.f
	950	8.3	207	14	T06753	T06753 EST04642 Fe
c	951	8.3	208	9	AV061018	AV061018 AV061018
	952	8.3	209	10	BB512967	BB512967 BB512967
	953	8.3	211	10	BB046643	BB046643 BB046643
	954	8.3	212	17	BH081780	BH081780 RPCI-24-2
c	955	8.3	213	9	AV059443	AV059443 AV059443

```
956       N74546      15   8.3        N74546 za06g05.s1
957       AI506131    15   8.3        AI506131 v130d07.x
c 958     AA268701    15   8.3        AA268701 va45g05.r
959       BG134898    15   8.3        BG134898 EST467790
960       AV380939    15   8.3        AV380939 AV380939
961       Z38990      15   8.3        Z38990 HSC0WA062 n
962       BI536686    15   8.3        BI536686 395994 MA
963       AV247158    15   8.3        AV247158 AV247158
964       AW619333    15   8.3        AW619333 707 MARC
965       AI615822    15   8.3        AI615822 v130d07.y
c 966     AI910015    15   8.3        AI910015 CM-BT233-
c 967     AA561402    15   8.3        AA561402 v130d07.r
968       AW336574    15   8.3        AW336574 22689 MAR
969       AV104793    15   8.3        AV104793 AV104793
970       BB593423    15   8.3        BB593423 BB593423
971       AA056378    15   8.3        AA056378 z165b01.s
c 972     AV006681    15   8.3        AV006681 AV006681
973       BG277295    15   8.3        BG277295 ux41h06.y
974       AL648155    15   8.3        AL648155 AL648155
975       AV053096    15   8.3        AV053096 AV053096
976       AV219997    15   8.3        AV219997 AV219997
977       AA867427    15   8.3        AA867427 vx98d12.r
978       AV051210    15   8.3        AV051210 AV051210
979       AW466240    15   8.3        AW466240 MU000075.
980       BI720974    15   8.3        BI720974 1031053C0
981       AA832311    15   8.3        AA832311 oc95g06.s
982       AV112405    15   8.3        AV112405 AV112405
983       BB148730    15   8.3        BB148730 BB148730
984       AV153877    15   8.3        AV153877 AV153877
985       BJ343736    15   8.3        BJ343736 BJ343736
986       AV208005    15   8.3        AV208005 AV208005
987       BB213055    15   8.3        BB213055 BB213055
988       BG162137    15   8.3        BG162137 dc46c01.y
989       BH612722    15   8.3        BH612722 SALK_0331
990       AV060103    15   8.3        AV060103 AV060103
991       AV123180    15   8.3        AV123180 AV123180
992       AV289211    15   8.3        AV289211 AV289211
c 993     AW224535    15   8.3        AW224535 EST302978
c 994     AW224536    15   8.3        AW224536 EST302979
995       BB175988    15   8.3        BB175988 BB175988
996       BF174751    15   8.3        BF174751 MYB3763 M
997       H08428      15   8.3        H08428 yl92h09.s1
c 998     AV013856    15   8.3        AV013856 AV013856
999       AV244904    15   8.3        AV244904 AV244904
c1000     BB034330    15   8.3        BB034330 BB034330
```

ALIGNMENTS

RESULT 1
AI962650
LOCUS   AI962650   410 bp   mRNA   linear   EST 08-MAR-2000
DEFINITION wq42e03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2473948 3'
            similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI962650
VERSION   AI962650.1 GI:5755363
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 410)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

---

Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 770   Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES   Location/Qualifiers
     source   1..410
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2473948"
                    /clone_lib="NCI_CGAP_GC6"
                    /tissue_type="pooled germ cell tumors"
                    /lab_host="DH10B"
                    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                    polylinker; Site_1: Not I, Site_2: Eco RI; Plasmid DNA
                    from the normalized library NCI_CGAP_GC4 was prepared, and
                    ss circles were made in vitro. Following HAP purification,
                    this DNA was used as tracer in a subtractive hybridization
                    reaction. The driver was PCR-amplified cDNAs from a pool
                    of 5,000 clones made from the same library (cloneIds
                    1257096-1258631, 1469064-1470983, and 1475592-1476743).
                    Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   120 a   93 c   93 g   104 t
ORIGIN

Query Match   83.4%;   Score 151;   DB 9;   Length 410;
Best Local Similarity 100.0%;   Pred. No. 1.7e-71;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   31 GACGTCTTCCGAAGGATTTTATTACTTATATGGACAATTGGGCCAGACAACACAGCT 90
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 GACGTCTTCCGAAGGATTTTATTACTTATATGGACAATTGGGCCAGACAACACAGCT 77

Qy   91 GAGCAAGAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTCTGTCATCCTGTAC 150
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   78 GAGCAAGAGAGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTCTGTCATCCTGTAC 137

Qy   151 CTCATGGTGATGATTGGAATGTTCTCTTTCA 181
           |||||||||||||||||||||||||||||||
Db   138 CTCATGGTGATGATTGGAATGTTCTCTTTCA 168

RESULT 2
AI654552
LOCUS   AI654552   429 bp   mRNA   linear   EST 17-DEC-1999
DEFINITION wb48b12.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2308895 3'
            similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
            CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION   AI654552
VERSION   AI654552.1 GI:4738531
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 429)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```
                   www-bio.llnl.gov/bbrp/image/image.html
                   Insert Length: 771   Std Error: 0.00
                   Seq primer: -40UP from Gibco
                   High quality sequence stop: 411.
FEATURES           Location/Qualifiers
     source        1..429
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:2308895"
                   /clone_lib="NCI_CGAP_GC6"
                   /tissue_type="pooled germ cell tumors"
                   /lab_host="DH10B"
                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                   polylinker; Site 1; Site 2: Eco RI; Plasmid DNA
                   from the normalized library NCI_CGAP_GC4 was prepared, and
                   ss circles were made in vitro. Following HAP purification,
                   this DNA was used as tracer in a subtractive hybridization
                   reaction. The driver was PCR-amplified cDNAs from a pool
                   of 5,000 clones made from the same library (cloneIDs
                   1257096-1258631, 1469064-1470983, and 1475592-1476743)."

                   Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT      127 a      100 c       97 g      104 t        1 others
ORIGIN

Query Match           83.4%;    Score 151;  DB 9;  Length 429;
Best Local Similarity 100.0%;  Pred. No. 1.7e-71;
Matches 151;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps    0;

Qy     31 GACGTCTTCCGAAGGATTTTTATTACTTATATGGACAATTGGCGCCCAGAACAACACAACAGCT 90
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GACGTCTTCCGAAGGATTTTTATTACTTATATGGACAATTGGCGCCCAGAACAACACAACAGCT 71

Qy     91 GAGCAAGAGGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTAC 150
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 GAGCAAGAGGGCCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTAC 131

Qy    151 CTCATGGTGATGATTGGAATGTTCTCTTTCA 181
          |||||||||||||||||||||||||||||||
Db    132 CTCATGGTGATGATTGGAATGTTCTCTTTCA 162


RESULT 3
AI246239
LOCUS        AI246239                372 bp    mRNA        linear   EST 28-JAN-1999
DEFINITION   qi29g04.x1 Soares NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1857942
             3' similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
             CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION    AI246239
VERSION      AI246239.1  GI:381636
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 372)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40UP from Gibco
             Insert Length: 921   Std Error: 0.00
             High quality sequence stop: 365.
FEATURES           Location/Qualifiers
     source        1..372
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:1857942"
                   /clone_lib="Soares NhHMPu_S1"
                   /tissue_type="Pooled human melanocyte, fetal heart, and
```

```
pregnant uterus"
                   /lab_host="DH10B"
                   /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                   (Pharmacia) with a modified polylinker; Site 1: Not I;
                   Site 2: Eco RI; Equal amounts of plasmid DNA from three
                   normalized libraries (melanocyte 2NbHM, pregnant uterus
                   NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                   were made in vitro. Following HAP purification, this DNA
                   was used as tracer in a subtractive hybridization
                   reaction. The driver was PCR-amplified cDNAs from pools of
                   5,000 clones made from the same 3 libraries. The pools
                   consisted of I.M.A.G.E. clones 260232-265223,
                   340488-345479, and 484488-489479."

BASE COUNT      115 a       89 c       86 g       82 t
ORIGIN

Query Match           79.0%;    Score 143;  DB 9;  Length 372;
Best Local Similarity 100.0%;  Pred. No. 3.9e-67;
Matches 143;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps    0;

Qy     39 CCGGAAGGATTTTTATTACTTATATGGACAATTGGCGCCCAGAACAACACAACAGCTGAGCAAGA 98
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCGGAAGGATTTTTATTACTTATATGGACAATTGGCGCCCAGAACAACACAACAGCTGAGCAAGA 60

Qy     99 GGGCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGT 158
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGCCTCCAAGCCAAAGTTGATGCTGAGAACTTCTACTATGTCATCCTGTACCTCATGGT 120

Qy    159 GATGATTGGAATGTTCTCTTTCA 181
          |||||||||||||||||||||||
Db    121 GATGATTGGAATGTTCTCTTTCA 143


RESULT 4
AI339609
LOCUS        AI339609                391 bp    mRNA        linear   EST 29-DEC-1998
DEFINITION   qq42a07.x1 Soares NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1935156
             3' similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
             CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION    AI339609
VERSION      AI339609.1  GI:4076536
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 391)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40UP from Gibco
             High quality sequence stop: 380.
FEATURES           Location/Qualifiers
     source        1..391
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:1935156"
                   /clone_lib="Soares NhHMPu_S1"
                   /tissue_type="Pooled human melanocyte, fetal heart, and
                   pregnant uterus"
                   /lab_host="DH10B"
                   /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                   (Pharmacia) with a modified polylinker; Site 1: Not I;
                   Site 2: Eco RI; Equal amounts of plasmid DNA from three
                   normalized libraries (melanocyte 2NbHM, pregnant uterus
                   NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                   were made in vitro. Following HAP purification, this DNA
                   was used as tracer in a subtractive hybridization
```

```
Db        557 AGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 501

RESULT 6
BG221966
LOCUS       BG221966                746 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST41783 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG221966
VERSION     BG221966.1  GI:13747987
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain.S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            , J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
  COMMENT   Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 547.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT    241 a    138 c    123 g    243 t    1 others
ORIGIN

Query Match    46.4%;  Score 84;  DB 12;  Length 746;
Best Local Similarity  100.0%;  Pred. No. 6.5e-35;
Matches  84;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532 ACTTTATCCAATTCACACAGACGCTGGAAGACGTCTTCCGAAGGATTTTATTACTTAT 591

Qy    61 ATGGACAATTGGCGCCAGAACACA 84
           ||||||||||||||||||||||||
Db   592 ATGGACAATTGGCGCCAGAACACA 615


RESULT 7
AW869303/c
LOCUS       AW869303                311 bp    mRNA    linear    EST 22-MAY-2000
DEFINITION  MR3-SN0067-240400-006-f11 SN0067 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW869303
VERSION     AW869303.1  GI:8003356
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 311)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
```

```
            reaction. The driver was PCR-amplified cDNAs from pools of
            5,000 clones made from the same 3 libraries. The pools
            consisted of I.M.A.G.E. clones 260232-265223,
            340488-345479, and 484488-489479."
BASE COUNT    119 a     93 c     93 g     86 t
ORIGIN

Query Match    65.7%;  Score 119;  DB 9;  Length 391;
Best Local Similarity  100.0%;  Pred. No. 4.8e-54;
Matches  119;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    63 GGACAATTGGCGCCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGC 122
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    34 GGACAATTGGCGCCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGC 93

Qy   123 TGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 181
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    94 TGAGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 152


RESULT 5
BG208163/c
LOCUS       BG208163                803 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST27654 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG208163
VERSION     BG208163.1  GI:13729850
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain.S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            , J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
  COMMENT   Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 550.
FEATURES             Location/Qualifiers
     source          1..803
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT    222 a    172 c    171 g    238 t
ORIGIN

Query Match    64.6%;  Score 117;  DB 12;  Length 803;
Best Local Similarity  100.0%;  Pred. No. 6.6e-53;
Matches  117;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    65 ACAATTGGCGCCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTG 124
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 ACAATTGGCGCCAGAACAACAGCTGAGCAAGAGGCCCTCCAAGCCAAAGTTGATGCTG 558

Qy   125 AGAACTTCTACTATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
             Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
             sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-SN0067-240
             400-006-f11&t3=2000-04-24&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 11
             High quality sequence stop: 74.
FEATURES             Location/Qualifiers
     source          1..311
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="SN0067"
                     /dev_stage="Adult"
                     /note="Organ: stomach normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT    60 a    80 c    76 g    95 t
ORIGIN

Query Match        29.3%;  Score 53;  DB 10;  Length 311;
Best Local Similarity 100.0%;  Pred. No. 4.6e-18;
Matches   53;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGACGATTTTTAT 53
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 ACTTTATCCAATTTCACACAGACGCTGGAAGACGTCTTCCGAAGACGATTTTTAT 124

RESULT 8
AW869303
LOCUS        AW869303      311 bp    mRNA      linear    EST 22-MAY-2000
DEFINITION   MR3-SN0067-240400-006-f11 SN0067 Homo sapiens cDNA, mRNA sequence.
ACCESSION    AW869303
VERSION      AW869303.1  GI:8003356
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 311)
  AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
  TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
  JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE    20202663
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
```

```
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-SN0067-240
             400-006-f11&t3=2000-04-24&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 11
             High quality sequence stop: 74.
FEATURES             Location/Qualifiers
     source          1..311
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="SN0067"
                     /dev_stage="Adult"
                     /note="Organ: stomach normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT    60 a    80 c    76 g    95 t
ORIGIN

Query Match        24.9%;  Score 45;  DB 10;  Length 311;
Best Local Similarity 100.0%;  Pred. No. 1.1e-13;
Matches   45;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    137 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 181
           |||||||||||||||||||||||||||||||||||||||||||||
Db     11 ATGTCATCCTGTACCTCATGGTGATGATTGGAATGTTCTCTTTCA 55

RESULT 9
AA935321
LOCUS        AA935321      121 bp    mRNA      linear    EST 07-JUL-1998
DEFINITION   oo71g09.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1571680 3'
             similar to SW:MINK_HUMAN P15382 ISK SLOW VOLTAGE-GATED POTASSIUM
             CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION    AA935321
VERSION      AA935321.1  GI:3092478
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 121)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html

             Trace considered overall poor quality
             Insert Length: 763   Std Error: 0.00
             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..121
                     /organism="Homo sapiens"
```

```
            /db_xref="taxon:9606"
            /clone="IMAGE:1571680"
            /clone_lib="NCI_CGAP_GC4"
            /lab_host="DH10B"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; 1st strand cDNA was prepared from 3 pooled
            germ cell tumors, and was then primed with a Not I -
            oligo(dT) primer. Double-stranded cDNA was ligated to Eco
            RI adaptors (Pharmacia), digested with Not I and cloned
            into the Not I and Eco RI sites of the modified pT7T3
            vector. Library is normalized. Library was constructed by
            Bento Soares and M. Fatima Bonaldo. "
BASE COUNT    33 a    31 c    28 g    29 t
ORIGIN

Query Match           19.3%;   Score 35;   DB 9;   Length 121;
Best Local Similarity  100.0%;   Pred. No. 2.6e-08;
Matches   35;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    147 GTACCTCATGGTGATGATTGATGGAATGTTCTCTTTCA 181
          ||||||||||||||||||||||||||||||||||||||
Db      1 GTACCTCATGGTGATGATTGATGGAATGTTCTCTTTCA 35

RESULT 10
BG938225
LOCUS       BG938225               351 bp    mRNA          linear   EST 11-JUN-2001
DEFINITION  1Abo14A12 Bovine Abomasum cDNA Library Bos taurus cDNA 5', mRNA
            sequence.
ACCESSION   BG938225
VERSION     BG938225.1   GI:14337597
KEYWORDS    EST.
SOURCE      cow.
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 351)
AUTHORS     Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.
TITLE       cDNA's from bovine abomasum tissue
JOURNAL     Unpublished (2001)
COMMENT     Contact: Dr. Stephen Moore
            Beef Genomics Laboratory
            Dept of AFNS, University of Alberta
            410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
            Tel: 780 492 0169
            Fax: 780 492 4265
            Email: smoore@afns.ualberta.ca
            The sequence best matches gb:AP001719 (Homo sapiens genomic DNA,
            chromosome 21q, section 63/105) in main database at high score of
            212.0 and E-value of 9e-53
            PCR PRimers
            FORWARD: M13 Forward
            BACKWARD: M13 Reverse
            Seq primer: T3 primer
            High quality sequence stop: 351
            POLYA=NO.
FEATURES             Location/Qualifiers
     source          1..351
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="Bovine Abomasum cDNA Library"
                     /sex="Two males and one female mixed"
                     /tissue_type="Gastrointestinal tissue (GIT)"
                     /cell_type="Epithelial"
                     /dev_stage="Young adult"
                     /lab_host="XL1-BlueMRF'-strain"
                     /note="Organ: Abomasum; Vector: Uni-ZAPXR; Site_1: EcoR
                     I; Site_2: Xho I"
BASE COUNT    104 a    89 c    76 g    82 t
ORIGIN


Query Match           16.6%;   Score 30;   DB 13;   Length 351;
Best Local Similarity  100.0%;   Pred. No. 1.6e-05;
Matches   30;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    115 GTTGATGCTGAGAACTTCTACTATGTCATC 144
          ||||||||||||||||||||||||||||||
Db    247 GTTGATGCTGAGAACTTCTACTATGTCATC 276

RESULT 11
D85797
LOCUS       D85797                 470 bp    mRNA          linear   EST 09-SEP-1996
DEFINITION  D85797 Rat 21 day old female ovary mRNA PMSG 3h Rattus norvegicus
            cDNA, mRNA sequence.
ACCESSION   D85797
VERSION     D85797.1   GI:1465812
KEYWORDS    EST.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 470)
AUTHORS     Miyamoto,K., Mizutani,T., Numata,Y. and Okada,T.
TITLE       Rat 21 day old female ovary mRNA
JOURNAL     Unpublished (1996)
COMMENT     Contact: Miyamoto,K.
            Institute for Molecular and Cellular Reg
            Gunma Univ. Biosignal Research Center
            Showa machi 3-39-15, Maebashi, Gunma 371, Japan
            Email: tmizu@sb.gunma-u.ac.jp
            5'-59.
FEATURES             Location/Qualifiers
     source          1..470
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /clone_lib="Rat 21 day old female ovary mRNA PMSG 3h"
                     /sex="female"
                     /tissue_type="ovary"
                     /dev_stage="21 day old"
                     /note="tissue_type=hypothalamus; dev_stage=adult rat"
                     4_others
BASE COUNT    116 a   127 c   128 g    95 t
ORIGIN


Query Match           14.4%;   Score 26;   DB 14;   Length 470;
Best Local Similarity  100.0%;   Pred. No. 0.0026;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    139 GTCATCCTGTACCTCATGGTGATGAT 164
          ||||||||||||||||||||||||||
Db    181 GTCATCCTGTACCTCATGGTGATGAT 206

RESULT 12
BG261965
LOCUS       BG261965              1003 bp    mRNA          linear   EST 13-FEB-2001
DEFINITION  602737384F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481325 5',
            mRNA sequence.
ACCESSION   BG261965
VERSION     BG261965.1   GI:12771781
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1003)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapps-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
```

```
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10316  row: g  column: 22
          High quality sequence stop: 535.
FEATURES             Location/Qualifiers
     source          1..1003
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4481325"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT    249 a    218 c    302 g    234 t
ORIGIN

Query Match      14.4%;   Score 26;  DB 12;  Length 1003;
Best Local Similarity  100.0%;  Pred. No. 0.0029;
Matches  26;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  139 GTCATCCTGTACCTCATGGTGATGAT 164
        ||||||||||||||||||||||||||
Db  235 GTCATCCTGTACCTCATGGTGATGAT 260

RESULT 13
AK008619
LOCUS       AK008619                1691 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male stomach cDNA, RIKEN full-length enriched
            library, clone:2200002I16:homolog to MINIMUM POTASSIUM ION
            CHANNEL-RELATED PEPTIDE 1 (MIRP1) (MINK-RELATED PEPTIDE 1), full
            insert sequence.
ACCESSION   AK008619
VERSION     AK008619.1  GI:12842913
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2200002I16.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99272253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,T., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 1691)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,K., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3'], cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES             Location/Qualifiers
     source          1..1691
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM DB:2200002I16"
                     /db_xref="MGD:MGI:1901204"
                     /db_xref="taxon:10090"
                     /clone="2200002I16"
                     /sex="male"
```

```
        CDS              /tissue_type="stomach"
                         /clone_lib="RIKEN full-length enriched mouse cDNA library"
                         /dev_stage="adult"
                         149..520
                         /note="data source:SPTR, source key:Q9WTW0, evidence:ISS
                         homolog to MINIMUM POTASSIUM ION CHANNEL-RELATED PEPTIDE 1
                         (MIRP1) (MINK-RELATED PEPTIDE 1)
                         putative"
                         /codon_start=1
                         /protein_id="BAB25781.1"
                         /db_xref="GI:12842914"
                         /db_xref="MGD:MGI:1916393"
                         /translation="MATLANLTQTLEDAFKKIFITYMDSWRRNTTAEEQALQARVDAE
                         NFYYVILVLMVMIGMPSFIVVAILVSTVKSKREHSQDPYHQYIVEDWQERYKSQILH
                         LEDSKATIHENMGATGFTVSP"
        polyA_signal     1673..1678
                         /note="putative"
        polyA_site       1691
                         /note="putative"
BASE COUNT     482 a    365 c    410 g    434 t
ORIGIN

Query Match             14.4%;  Score 26;  DB 11;  Length 1691;
Best Local Similarity   100.0%;  Pred. No. 0.0031;
Matches    26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    139 GTCATCCTGTACCTCATGGTGATGAT 164
          ||||||||||||||||||||||||||
Db    293 GTCATCCTGTACCTCATGGTGATGAT 318

RESULT 14
BB564873
LOCUS       BB564873      314 bp    mRNA      linear    EST 29-NOV-2000
DEFINITION  BB564873 RIKEN full-length enriched, adult male stomach Mus
            musculus cDNA clone 2200002116 5', mRNA sequence.
ACCESSION   BB564873
VERSION     BB564873.1  GI:11455765
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 314)
AUTHORS     Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.
            , Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodoyama,Y.,
            Imotani K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno
            ,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,
            Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C.,
            Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A.,
            Shiraki,K., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka
            ,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshiki,A., Muramatsu.M. and Hayashizaki,Y.
TITLE       RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
JOURNAL     Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci.P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh,M., Kitsuna,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
            Automated filtration-based high-throughput plasmid preparation
```

```
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci.P. and Hayashizaki,Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1..314
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="2200002116"
                     /clone_lib="RIKEN full-length enriched, adult male
                     stomach"
                     /sex="male"
                     /tissue_type="stomach"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /note="Site_1: XhoI; Site_2: SstI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGCGGCCGCCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA
                     was cleaved with XhoI and SstI."
BASE COUNT       98 a      77 c      74 g      65 t
ORIGIN

Query Match             12.2%;  Score 22;  DB 10;  Length 314;
Best Local Similarity   100.0%;  Pred. No. 0.37;
Matches    22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 ATTTTTATTACTTATATGGACA 67
          ||||||||||||||||||||||
Db    200 ATTTTTATTACTTATATGGACA 221

RESULT 15
H77654
LOCUS       H77654        365 bp    mRNA      linear    EST 09-NOV-1995
DEFINITION  yu24d09.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:234737 3', mRNA sequence.
ACCESSION   H77654
VERSION     H77654.1  GI:1055743
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 365)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 968
            High quality sequence stops: 281
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
              Insert Length: 968   Std Error: 0.00
              Seq primer: Promeg -21m13
              High quality sequence stop: 281.
FEATURES      Location/Qualifiers
     source   1..365
              /organism="Homo sapiens"
              /db_xref="GDB:3787914"
              /db_xref="taxon:9606"
              /clone="IMAGE:234737"
              /sex="male"
              /dev_stage="20 week-post conception fetus"
              /lab_host="DH10B (ampicillin resistant)"
              /clone_lib="Soares fetal liver spleen 1NFLS"
              /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
              with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
              1st strand cDNA was primed with a Pac I - oligo(dT) primer
              [5' AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
              double-stranded cDNA was ligated to Eco RI adaptors
              (Pharmacia), digested with Pac I and cloned into the Pac I
              and Eco RI sites of the modified pT7T3 vector. Library
              went through one round of normalization. Library
              constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT    105 a     71 c     74 g    109 t
ORIGIN

Query Match     11.6%;   Score 21;  DB 14;  Length 365;
Best Local Similarity 100.0%;   Pred. No. 1.3;
Matches  21; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy    124 GAGAACTTCTACTATGTCATC 144
              ||||||||||||||||||||
Db    322 GAGAACTTCTACTATGTCATC 342


RESULT 16
BH827709/c
LOCUS         BH827709        287 bp    DNA     linear  GSS 20-MAY-2002
DEFINITION    BACPP26-G18.y Pristionchus pacificus BAC ends Pristionchus
              pacificus genomic, DNA sequence.
ACCESSION     BH827709
VERSION       BH827709.1  GI:21021448
KEYWORDS      GSS.
SOURCE        Pristionchus pacificus
  ORGANISM    Pristionchus pacificus
              Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
              Neodiplogasteridae; Pristionchus.
REFERENCE     1  (bases 1 to 287)
  AUTHORS     Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R., Raddatz
              ,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A., Jesse,T.,
              Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
  TITLE       A BAC-based genetic linkage map of the nematode Pristionchus
              pacificus
  JOURNAL     Unpublished (2002)
COMMENT       Contact: Sommer RJ
              Evolutionary Biology
              Max-Planck-Institute for Developmental Biology
              Spemannstr. 37-39, Tuebingen D-72076, Germany
              Tel: 0049707160 1371
              Fax: 0049707160 1498
              Email: ralf.sommer@tuebingen.mpg.de
              Class: BAC ends.
FEATURES      Location/Qualifiers
     source   1..287
              /organism="Pristionchus pacificus"
              /strain="var. California"
              /db_xref="taxon:54126"
              /clone_lib="Pristionchus pacificus BAC ends"
BASE COUNT     83 a     60 c     78 g     66 t
ORIGIN

Query Match     10.5%;   Score 19;  DB 17;  Length 287;
Best Local Similarity 100.0%;   Pred. No. 16;
Matches  19; Conservative  0; Mismatches  0; Indels  0; Gaps  0;
```

```
Qy     27 GGAAGACGTCTTCCGAAGG 45
             ||||||||||||||||||
Db    209 GGAAGACGTCTTCCGAAGG 191


RESULT 17
BH822585/c
LOCUS         BH822585        341 bp    DNA     linear  GSS 20-MAY-2002
DEFINITION    BACPP17-F14.y Pristionchus pacificus BAC ends Pristionchus
              pacificus genomic, DNA sequence.
ACCESSION     BH822585
VERSION       BH822585.1  GI:20999456
KEYWORDS      GSS.
SOURCE        Pristionchus pacificus
  ORGANISM    Pristionchus pacificus
              Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
              Neodiplogasteridae; Pristionchus.
REFERENCE     1  (bases 1 to 341)
  AUTHORS     Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R., Raddatz
              ,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A., Jesse,T.,
              Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
  TITLE       A BAC-based genetic linkage map of the nematode Pristionchus
              pacificus
  JOURNAL     Unpublished (2002)
COMMENT       Contact: Sommer RJ
              Evolutionary Biology
              Max-Planck-Institute for Developmental Biology
              Spemannstr. 37-39, Tuebingen D-72076, Germany
              Tel: 0049707160 1371
              Fax: 0049707160 1498
              Email: ralf.sommer@tuebingen.mpg.de
              Class: BAC ends.
FEATURES      Location/Qualifiers
     source   1..341
              /organism="Pristionchus pacificus"
              /strain="var. California"
              /db_xref="taxon:54126"
              /clone_lib="Pristionchus pacificus BAC ends"
BASE COUNT     96 a     66 c     97 g     82 t
ORIGIN

Query Match     10.5%;   Score 19;  DB 17;  Length 341;
Best Local Similarity 100.0%;   Pred. No. 16;
Matches  19; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy     27 GGAAGACGTCTTCCGAAGG 45
             ||||||||||||||||||
Db    280 GGAAGACGTCTTCCGAAGG 262


RESULT 18
BH292512
LOCUS         CH230-44B10.TV CHORI-230  539 bp    DNA     linear  GSS 30-NOV-2001
DEFINITION    CH230-44B10.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
              CH230-44B10, DNA sequence.
ACCESSION     BH292512
VERSION       BH292512.1  GI:17204920
KEYWORDS      GSS.
SOURCE        Norway rat.
  ORGANISM    Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE     1  (bases 1 to 539)
  AUTHORS     Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
              ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.M.
  TITLE       Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL     Unpublished (1999)
COMMENT       Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
```

```
                        9712 Medical Center Dr., Rockville, MD 20850, USA
                        Tel: 301 838 0200
                        Fax: 301 838 0208
                        Email: szhao@tigr.org
                        Clones are derived from the rat BAC library CHORI-230
                        (http://www.chori.org/bacpac/rat230.htm). For BAC library
                        availability, please contact Pieter de Jong (pdejong@mail.cho.org).
                        Clones may be purchased from BACPAC Resources
                        (http://www.chori.org/bacpac/or ering_information.htm). BAC end
                        page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
                        Plate: 44 row: E column: 10
                        Seq primer: T7
                        Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .539
                        /organism="Rattus norvegicus"
                        /strain="BN/SsNHsd/MCW"
                        /db_xref="taxon:10116"
                        /clone="CH230-44E10"
                        /sex="Female"
                        /cell_type="Brain"
                        /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                        CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                        Pieter de Jong"
BASE COUNT     179 a     70 c     91 g    199 t
ORIGIN

Query Match           10.5%;  Score 19;  DB 17;  Length 539;
Best Local Similarity 100.0%;  Pred. No. 17;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    41 GAAGGATTTTTATTACTTA 59
         ||||||||||||||||||||
Db   193 GAAGGATTTTTATTACTTA 211

RESULT 19
BH302149
BH302149
LOCUS       BH302149     799 bp    DNA    linear   GSS 30-NOV-2001
DEFINITION  CH230-85A14.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-85A14, DNA sequence.
ACCESSION   BH302149
VERSION     BH302149.1  GI:17214557
KEYWORDS    GSS.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 799)
AUTHORS     Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
TITLE       Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL     Unpublished (1999)
COMMENT     Other_GSSs: CH230-85A14.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pinter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 85 row: A column: 14
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .799
                        /organism="Rattus norvegicus"
                        /strain="BN/SsNHsd/MCW"
                        /db_xref="taxon:10116"
                        /clone="CH230-85A14"
                        /clone_lib="CHORI-230 Segment 1"
                        /sex="Female"
                        /cell_type="Brain"
                        /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                        CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                        Pieter de Jong"
BASE COUNT     247 a    115 c    131 g    306 t
ORIGIN

Query Match           10.5%;  Score 19;  DB 17;  Length 799;
Best Local Similarity 100.0%;  Pred. No. 18;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    41 GAAGGATTTTTATTACTTA 59
         ||||||||||||||||||||
Db   180 GAAGGATTTTTATTACTTA 198

RESULT 20
AL557074
AL557074
LOCUS       AL557074     938 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  LTI_FL012_TC1 Homo sapiens cDNA clone CSODH005YO23 5 prime
            , mRNA sequence.
ACCESSION   AL557074
VERSION     AL557074.1  GI:12900327
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 938)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .938
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="CSODH005YO23"
                        /clone_lib="LTI_FL012_TC1"
                        /tissue_type="T cells from T cell leukemia"
                        /lab_host="DH10B"
                        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                        with a Noti-oligo(dT) primer. Five prime end enriched.
                        double-stranded cDNA was digested with Not I and cloned
                        into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
                        Library was constructed by Life Technologies. Contact :
                        Feng Liang Life Technologies, a division of Invitrogen
                        9800 Medical Center Drive Rockville, Maryland 20850, USA
                        Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
                        http://fulllength.invitrogen.com"
BASE COUNT     169 a    295 c    319 g    153 t
ORIGIN

Query Match           10.5%;  Score 19;  DB 9;  Length 938;
Best Local Similarity 100.0%;  Pred. No. 19;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    97 GAGGCCCTCCAAGCCAAAG 115
         |||||||||||||||||||
Db   816 GAGGCCCTCCAAGCCAAAG 834
```

```
RESULT 21
BH644242
LOCUS       BH644242        139 bp    DNA    linear    GSS 14-FEB-2002
DEFINITION  1008056H01.x1 1008 - RescueMu Grid I Zea mays genomic, DNA
            sequence.
ACCESSION   BH644242
VERSION     BH644242.1  GI:18673039
KEYWORDS    GSS.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 139)
AUTHORS     Walbot,V.
TITLE       Maize genomic sequences found using engineered RescueMu transposon
JOURNAL     Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site so sequence was trimmed. Post-ligation
            sequence submitted separately.
            Plate: 1008056 row: 5
            Class: transposon-tagged.
FEATURES    Location/Qualifiers
     source 1..139
            /organism="Zea mays"
            /cultivar="mixed background W23/A188/B73"
            /db_xref="taxon:4577"
            /clone_lib="1008 - RescueMu Grid I"
            /tissue_type="leaf"
            /dev_stage="adult"
            /lab_host="DH10B"
            /note="Organ: leaf; Vector: RescueMu (engineered from
            pBlueScript backbone); Site 1: BamHI; Site 2: BglII;
            RescueMu is a 4.9 kb, modified maize Mu transposon
            designed to allow plasmid rescue from total genomic DNA.
            Mu elements insert preferentially into transcription
            units. For more information on RescueMu, go to the web
            site www.zmdb.iastate.edu and follow the links for
            'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
            extracted from leaf punches, double digested using BamHI
            and BglII, and ligated to form circular plasmids. DH10B
            cells were transformed and then screened on LB plates with
            ampicillin."
BASE COUNT    25 a    51 c    32 g    31 t
ORIGIN

Query Match   9.9%;  Score 18;  DB 17;  Length 139;
Best Local Similarity  100.0%;  Pred. No. 50;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     23 CGCTGGAAGACGTCTTCC 40
          ||||||||||||||||||
Db     87 CGCTGGAAGACGTCTTCC 104

RESULT 22
BM416858/c
LOCUS       BM416858        169 bp    mRNA    linear    EST 28-JAN-2002
DEFINITION  953004F12.x2 953 - Immature ear with common ESTs screened by
            Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION   BM416858
VERSION     BM416858.1  GI:18383658
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 169)
AUTHORS     Walbot,V.
TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
            University
JOURNAL     Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 953004 row: F column: 12.
FEATURES    Location/Qualifiers
     source 1..169
            /organism="Zea mays"
            /cultivar="OH43"
            /db_xref="taxon:4577"
            /clone_lib="953 - Immature ear with common ESTs screened
            by Schmidt lab"
            /tissue_type="Inflorescence meristem - floral organ
            primordia"
            /dev_stage="0.5 cm to 2 cm"
            /lab_host="Stratagene XLOLR"
            /note="Organ: Immature ear; Vector: ZAP Express (pBK-CMV);
            Site 1: EcoRI; Site 2: XhoI; RNA from library 606 was
            filtered for common ESTs found in 606."
BASE COUNT    30 a    44 c    42 g    53 t
ORIGIN

Query Match   9.9%;  Score 18;  DB 13;  Length 169;
Best Local Similarity  100.0%;  Pred. No. 51;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     22 ACGCTGGAAGACGTCTTC 39
          ||||||||||||||||||
Db    132 ACGCTGGAAGACGTCTTC 115

RESULT 23
AV358849/c
LOCUS       AV358849        251 bp    mRNA    linear    EST 13-NOV-1999
DEFINITION  AV358849 RIKEN full-length enriched, in vitro fertilized eggs Mus
            musculus cDNA clone 7420430N10 3', mRNA sequence.
ACCESSION   AV358849
VERSION     AV358849.1  GI:6405851
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 251)
AUTHORS     Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
            Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
            Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
            ,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
            Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
            Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
            Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y.,
            Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
            Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE       RIKEN Mouse ESTs (Konno,H. et al 1999)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp.
```

URL:http://genome.gsc.riken.go.jp/
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki ,Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,K., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Carninci,P. and Hayashizaki,Y.
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
source　Location/Qualifiers
　　　　1..251
　　　　/organism="Mus musculus"
　　　　/strain="C57BL/6J"
　　　　/db_xref="taxon:10090"
　　　　/clone="742043N10"
　　　　/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"
　　　　/sex="female"
　　　　/tissue_type="in vitro fertilized eggs"
　　　　/dev_stage="egg"
　　　　/lab_host="DH10B"
　　　　/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAAATCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT　71 a　53 c　40 g　87 t
ORIGIN

Query Match　9.9%;　Score 18;　DB 10;　Length 251;
Best Local Similarity　100.0%;　Pred. No. 54;
Matches　18;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　142 ATCCTGTACCTCATGGTG 159
　　　　　||||||||||||||||||
Db　97 ATCCTGTACCTCATGGTG 80

RESULT 24
AW921614/c
LOCUS　AW921614　401 bp　mRNA　linear　EST 25-MAY-2000
DEFINITION　EST352918 Rat gene index, normalized rat, norvegicus, Bento Soares Rattus norvegicus cDNA clone RGIIB80 5' end, mRNA sequence.
ACCESSION　AW921614
VERSION　AW921614.1　GI:8087438
KEYWORDS　EST.
SOURCE　Norway rat.
ORGANISM　Rattus norvegicus
　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE　1 (bases 1 to 401)
AUTHORS　Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
TITLE　Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat

JOURNAL　Gene Index
　Unpublished (1998)
COMMENT　Contact: Lee, NH
　The Institute for Genomic Research
　9712, Medical Center Drive, Rockville, MD 20850, USA
　Tel: (301)-838-3529
　Fax: (301)-838-0208
　Email: nhlee@tigr.org
　This clone is available through the ATCC, contact the ATCC tel#703-365-2700 for further information
　Seq primer: M13 Reverse.
FEATURES
source　Location/Qualifiers
　　　　1..401
　　　　/organism="Rattus norvegicus"
　　　　/db_xref="taxon:10016"
　　　　/clone="RGIIB80"
　　　　/clone_lib="Rat gene index, normalized rat, norvegicus, Bento Soares"
　　　　/tissue_type="mix - brain, ovary, placenta, kidney, lung, liver, embryo, heart, muscle, spleen"
　　　　/lab_host="SOLR"
　　　　/note="Vector: pBlueScript SK(-); Site_1: EcoRl; Site_2: XhoI; Estimated insert size approx.1 kB"

BASE COUNT　84 a　101 c　109 g　107 t
ORIGIN

Query Match　9.9%;　Score 18;　DB 10;　Length 401;
Best Local Similarity　100.0%;　Pred. No. 58;
Matches　18;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　81 CACAACAGCTGAGCAAGA 98
　　　　||||||||||||||||||
Db　184 CACAACAGCTGAGCAAGA 167

RESULT 25
AA984762
LOCUS　AA984762　414 bp　mRNA　linear　EST 27-MAY-1998
DEFINITION　am90g02.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone IMAGE:1630418 3', mRNA sequence.
ACCESSION　AA984762
VERSION　AA984762.1　GI:3163287
KEYWORDS　EST.
SOURCE　human.
ORGANISM　Homo sapiens
　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1 (bases 1 to 414)
AUTHORS　Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE　WashU-NCI human EST Project
JOURNAL　Unpublished (1997)
COMMENT　Contact: Wilson RK
　Washington University School of Medicine
　4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
　Tel: 314 286 1800
　Fax: 314 286 1810
　Email: est@watson.wustl.edu
　This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
　Possible reversed clone: polyT not found
　Seq primer: -40m13 fwd. ET from Amersham
　High quality sequence stop: 414.
FEATURES
source　Location/Qualifiers
　　　　1..414
　　　　/organism="Homo sapiens"
　　　　/db_xref="taxon:9606"
　　　　/clone="IMAGE:1630418"
　　　　/clone_lib="Stratagene schizo brain S11"
　　　　/sex="male"
　　　　/tissue_type="schizophrenic brain S-11 frontal lobe"

```
            /dev_stage="34 years old"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Vector: Bluescript SK-; Site_1: EcoRI; Library
            constructed from S-11 frontal lobe, male, 34 years old,
            50% caucasian, 50% Aleutian. Schizophrenic suicide.
            Random primed into EcoRI site of ZAP II Vector. Mass
            excised. Custom library. Avg insert length 1.4kb.
            Material obtained by Johnston N., Torrey, E.F., Yolken R.,
            and the Stanley Neuropathology Consortium - Analysis of
            RNAs from the Brains of Individuals with Psychiatric
            Diseases (Unpublished) Stanley Neurovirology Laboratory,
            Johns Hopkins School of Medicine, Baltimore MD."
BASE COUNT    143 a    66 c    66 g    119 t
ORIGIN

Query Match    9.9%;  Score 18;  DB 9;  Length 414;
Best Local Similarity 100.0%;  Pred. No. 59;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   46 ATTTTTATTACTTATATG 63
        |||||||||||||||||
Db   70 ATTTTTATTACTTATATG 87

RESULT 26
BE478577
LOCUS       BE478577   419 bp   mRNA   linear   EST 28-AUG-2000
DEFINITION  162895 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE478577
VERSION     BE478577.1  GI:9598110
KEYWORDS    EST.
SOURCE      cow.
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 419)
AUTHORS     Sonstegard,T.S., Capuco,A.V., Van Tassell,C.P., Ashwell,M.S. and
            Wells,K.D.
TITLE       Mapping of Expressed Sequence Tags from a normalized bovine mammary
            gland cDNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Sonstegard TS
            USDA, ARS, Beltsville Agricultural Research Center
            Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
            Tel: 301 504 8416
            Fax: 301 504 8414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR Primers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 21  row: N  column: 7
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES    Location/Qualifiers
   source   1..419
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            /clone_lib="BARC 5BOV"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
            Library made from pooled mRNA isolated from mammary
            tissues at eight physiological, developmental, and disease
            states."
BASE COUNT    118 a    77 c    112 g    112 t
ORIGIN

Query Match    9.9%;  Score 18;  DB 10;  Length 419;
Best Local Similarity 100.0%;  Pred. No. 59;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   53 TTACTTATATGGACAATT 70
        |||||||||||||||||
Db   45 TTACTTATATGGACAATT 62

RESULT 27
AL040827/c
LOCUS       AL040827   425 bp   mRNA   linear   EST 29-FEB-2000
DEFINITION  DKFZp434F1615_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434F1615 5', mRNA sequence.
ACCESSION   AL040827
VERSION     AL040827.1  GI:5409772
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 425)
AUTHORS     Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE       EST (Bloecker, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp434F1615) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES    Location/Qualifiers
   source   1..425
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="DKFZp434F1615"
            /clone_lib="434 (synonym: htes3)"
            /tissue_type="testis"
            /dev_stage="adult"
            /lab_host="DH10B"
            /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT    200 a    45 c    60 g    120 t
ORIGIN

Query Match    9.9%;  Score 18;  DB 9;  Length 425;
Best Local Similarity 100.0%;  Pred. No. 59;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   45 GATTTTTTATTACTTATAT 62
        ||||||||||||||||||
Db  318 GATTTTTTATTACTTATAT 301

RESULT 28
BE941069
LOCUS       BE941069   471 bp   mRNA   linear   EST 03-OCT-2000
DEFINITION  EST420648 MGHG Medicago truncatula cDNA clone pMGHG-2B24, mRNA
            sequence.
ACCESSION   BE941069
VERSION     BE941069.1  GI:10518828
KEYWORDS    EST.
SOURCE      barrel medic.
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 471)
```

```
AUTHORS     Cote,F., Ojanen-Reuths,T., Hahn,M.G., VandenBosch,K., Hur,J.,                    Tel: 314 286 1800
            Beremand,P., Endre,G., Town,C.D., Bowman,C.L., Craven,M.B. and Cho               Fax: 314 286 1810
            ,J., Fraser,C.M.                                                                 Email: est@watson.wustl.edu
TITLE       ESTs from seedling roots of Medicago truncatula after treatment                 The library was constructed by Brandi Chiapelli and Dr. James
            with beta glucan elicitor preparation from Phytophthora sojae                    McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
JOURNAL     Unpublished (2000)                                                               Washington University, St. Louis. DNA Sequencing by: Washington
COMMENT     Contact: Michael G. Hahn                                                         University Genome Sequencing Center St. Louis.
            Complex Carbohydrate Research Center                                             Seq primer: -40RP from Gibco
            University of Georgia                                                            High quality sequence stop: 381.
            220 Riverbend Road, Athens, GA 30602-4712, USA                          FEATURES             Location/Qualifiers
            Tel: 706-542-4457                                                                source          1..479
            Fax: 706-542-4412                                                                                /organism="Parastrongyloides trichosuri"
            Email: hahn@ccrc.uga.edu                                                                         /db_xref="taxon:13110"
            University of Georgia name: G268921e TIGR sequence name: MTJAD12TK                               /clone_lib="Parastrongyloides trichosuri IL pAMP1 v1
            More information is available at: http://chrysie.tamu.edu/medicago                               Chiapelli McCarter"
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).                                                      /dev_stage="Infective Larvae"
FEATURES             Location/Qualifiers                                                                     /lab_host="DH10B"
            source          1..471                                                                           /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                            /organism="Medicago truncatula"                                                  The library was constructed by Brandi Chiapelli and Dr.
                            /cultivar="A17"                                                                  James McCarter at Washington University, St. Louis. The
                            /db_xref="taxon:3880"                                                            cDNA was made by using Dynabead oligo-dT priming (Dynal).
                            /clone="pMGHG-2B24"                                                              PCR based library using a modified protocol from the
                            /clone_lib="MGHG"                                                                SMART PCR cDNA Synthesis Kit from Clontech. Directionally
                            /tissue_type="Roots from four day old seedlings"                                cloned into the UDG sites of pAMP1. Nematodes were
                            /dev_stage="2 days after treatment with beta glucan                              provided by Dr. Warwick Grant of AgRearch, New Zealand
                            elicitor preparation from Phytophthora sojae"                                    (warwick.grant@agresearch.co.nz)."
                            /lab_host="E. coli strain XLOLR"                         BASE COUNT     142 a         64 c         78 g        195 t
                            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:  ORIGIN
                            XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                            was directionally ligated into the Unizap XR vector from Query Match        9.9%; Score 18;  DB 13;  Length 479;
                            Stratagene and packaged using Gigapack III Gold packaging  Best Local Similarity   100.0%;  Pred. No. 60;
                            extracts. Plasmids containing cDNA inserts were excised    Matches 18; Conservative  0; Mismatches  0; Indels   0; Gaps  0;
                            from the recombinant lambda-Zap phage using Ex-assist
                            helper phage and propagated in XLOLR cells."            Qy     48 TTTTATTACTTATATGGA 65
BASE COUNT     135 a        100 c         73 g       163 t                                    |||||||||||||||||||
ORIGIN                                                                              Db     31 TTTTATTACTTATATGGA 48

Query Match       9.9%;  Score 18;  DB 12;  Length 471;                            RESULT 30
Best Local Similarity   100.0%;  Pred. No. 60;                                     AI649434
Matches 18; Conservative  0; Mismatches  0; Indels   0; Gaps  0;                   LOCUS       AI649434         580 bp    mRNA       linear     EST 30-APR-1999
                                                                                   DEFINITION  603004D02.x2 603 - stressed root cDNA library from Wang/Bohnert lab
Qy     45 GATTTTTATTACTTATAT 62                                                                 Zea mays cDNA, mRNA sequence.
          |||||||||||||||||||                                                      ACCESSION   AI649434
Db     13 GATTTTTATTACTTATAT 30                                                    VERSION     AI649434.1  GI:4730268
                                                                                   KEYWORDS    EST.
RESULT 29                                                                          SOURCE      Zea mays.
BI742594                                                                           ORGANISM    Zea mays
LOCUS       BI742594         479 bp    mRNA       linear     EST 25-SEP-2001                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DEFINITION  kx33a05.y1 Parastrongyloides trichosuri IL pAMP1 v1 Chiapelli                     Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            McCarter Parastrongyloides trichosuri cDNA 5' similar to WP:CE24272                clade; Panicoideae; Andropogoneae; Zea.
            Y41G9A.4 glutamate receptor ;, mRNA sequence.                          REFERENCE   1 (bases 1 to 580)
ACCESSION   BI742594                                                               AUTHORS     Walbot.V.
VERSION     BI742594.1  GI:15764396                                                TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
KEYWORDS    EST.                                                                                University
SOURCE      Parastrongyloides trichosuri.                                          JOURNAL     Unpublished (1999)
ORGANISM    Parastrongyloides trichosuri                                          COMMENT     Contact: Walbot V
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;                             Department of Biological Sciences
            Panagrolaimoidea; Strongyloididae; Parastrongyloides.                              Stanford University
REFERENCE   1 (bases 1 to 479)                                                                 855 California Ave, Palo Alto, CA 94304, USA
AUTHORS     McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.                Tel: 650 723 2227
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.                Fax: 650 725 8221
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.               Email: walbot@stanford.edu
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,L., Steptoe                 Plate: 603004 row: D column: 02.
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.   FEATURES             Location/Qualifiers
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and                   source          1..580
            Wilson,R.                                                                                        /organism="Zea mays"
TITLE       The Washington Univ. Nematode EST Project, 1999                                                  /cultivar="B73"
JOURNAL     Unpublished (1999)                                                                               /db_xref="taxon:4577"
COMMENT     Contact: McCarter JP                                                                             /clone_lib="603 - stressed root cDNA library from
            The Washington Univ. Nematode EST Project, 1999                                                  Wang/Bohnert lab"
            Washington University School of Medicine                                                         /tissue_type="seedling"
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA                                     /dev_stage="salt stress"
```

TITLE   Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 690)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   On Feb 5, 1999 this sequence version replaced gi:2754668.
            AG003766: Submitted (07-Jan-1998).
FEATURES   Location/Qualifiers
     source 1..690
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T485XN"
BASE COUNT 153 a   129 c   140 t   7 others
ORIGIN

Query Match   9.9%;   Score 18; DB 17; Length 690;
Best Local Similarity   100.0%; Pred. No. 63;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   78 GAACACAACAGCTGAGCA 95
         |||||||||||||||||
Db   81 GAACACAACAGCTGAGCA 98

RESULT 33
AG009958
LOCUS   AG009958   692 bp   DNA   linear   GSS 14-APR-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
            sequence.
ACCESSION   AG009958 AG003769
VERSION   AG009958.1 GI:3289944
KEYWORDS   GSS.
SOURCE   Homo sapiens DNA, clone:T485XN.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 692)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE   Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 692)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   On Feb 5, 1999 this sequence version replaced gi:2754671.
            AG003769: Submitted (07-Jan-1998).
FEATURES   Location/Qualifiers
     source 1..692
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T485XN"
BASE COUNT 187 a   184 c   136 g   182 t   3 others
ORIGIN

Query Match   9.9%;   Score 18; DB 17; Length 692;
Best Local Similarity   100.0%; Pred. No. 63;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   78 GAACACAACAGCTGAGCA 95
         |||||||||||||||||
Db   664 GAACACAACAGCTGAGCA 681

            /lab_host="E. coli XL Gold"
            /note="Organ: root; Vector: pBluescriptII SK(+) XR;
            Seedling stressed root cDNA library from Wang/Bohnert lab"
BASE COUNT 115 a   164 c   171 g   130 t
ORIGIN

Query Match   9.9%;   Score 18; DB 9; Length 580;
Best Local Similarity   100.0%; Pred. No. 62;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   23 CGCTGGAAGACGTCTTCC 40
         ||||||||||||||||||
Db   427 CGCTGGAAGACGTCTTCC 444

RESULT 31
AG009962/c
LOCUS   AG009962   682 bp   DNA   linear   GSS 14-APR-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
            sequence.
ACCESSION   AG009962 AG003773
VERSION   AG009962.1 GI:3289948
KEYWORDS   GSS.
SOURCE   Homo sapiens DNA, clone:T485XN.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 682)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE   Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 682)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   On Feb 5, 1999 this sequence version replaced gi:2754675.
            AG003773: Submitted (07-Jan-1998).
FEATURES   Location/Qualifiers
     source 1..682
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T485XN"
BASE COUNT 153 a   102 c   218 g   207 t   2 others
ORIGIN

Query Match   9.9%;   Score 18; DB 17; Length 682;
Best Local Similarity   100.0%; Pred. No. 63;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   78 GAACACAACAGCTGAGCA 95
         ||||||||||||||||||
Db   219 GAACACAACAGCTGAGCA 202

RESULT 32
AG009955
LOCUS   AG009955   690 bp   DNA   linear   GSS 14-APR-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
            sequence.
ACCESSION   AG009955 AG003766
VERSION   AG009955.1 GI:3289941
KEYWORDS   GSS.
SOURCE   Homo sapiens DNA, clone:T485XN.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 690)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.

RESULT 34
AG009957
LOCUS   AG009957   700 bp   DNA   linear   GSS 14-APR-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone:T485XN, genomic survey
            sequence.
ACCESSION   AG009957 AG003768
VERSION   AG009957.1 GI:3289943
KEYWORDS   GSS.
SOURCE   Homo sapiens DNA, clone:T485XN.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 700)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 700)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   On Feb 5, 1999 this sequence version replaced gi:2754670.
            AG003768: Submitted (07-Jan-1998).
FEATURES   Location/Qualifiers
     source   1..700
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="T485XN"
BASE COUNT   193 a   203 c   127 g   171 t
ORIGIN

Query Match   9.9%; Score 18; DB 17; Length 700;
Best Local Similarity 100.0%; Pred. No. 63;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;   6 others

Qy   78 GAACACAACAGCTGAGCA 95
              ||||||||||||||||||
Db   572 GAACACAACAGCTGAGCA 589

RESULT 35
AG009959
LOCUS   AG009959   702 bp   DNA   linear   GSS 14-APR-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone:T485XN, genomic survey
            sequence.
ACCESSION   AG009959 AG003770
VERSION   AG009959.1 GI:3289945
KEYWORDS   GSS.
SOURCE   Homo sapiens DNA, clone:T485XN.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 702)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1998)
REFERENCE   2 (bases 1 to 702)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
            Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
            Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
            Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT   On Feb 5, 1999 this sequence version replaced gi:2754672.
            AG003770: Submitted (07-Jan-1998).
FEATURES   Location/Qualifiers
     source   1..702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="T485XN"
BASE COUNT   190 a   193 c   136 g   179 t
ORIGIN

Query Match   9.9%; Score 18; DB 17; Length 702;
Best Local Similarity 100.0%; Pred. No. 64;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;   4 others

Qy   78 GAACACAACAGCTGAGCA 95
              ||||||||||||||||||
Db   668 GAACACAACAGCTGAGCA 685

RESULT 36
BH556024/c
LOCUS   BH556024   704 bp   DNA   linear   GSS 14-DEC-2001
DEFINITION BOGWT27TF BOGW Brassica oleracea genomic clone BOGWT27, DNA
            sequence.
ACCESSION   BH556024
VERSION   BH556024.1 GI:17807804
KEYWORDS   GSS.
SOURCE   Brassica oleracea.
  ORGANISM Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 704)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE   Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
            Other_GSSs: BOGWT27TR
COMMENT   Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..704
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGWT27"
                     /clone_lib="BOGW"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT   131 a   180 c   139 g   254 t
ORIGIN

Query Match   9.9%; Score 18; DB 17; Length 704;
Best Local Similarity 100.0%; Pred. No. 64;
  Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   112 AAAGTTGATGCTGAGAAC 129
              ||||||||||||||||||
Db   389 AAAGTTGATGCTGAGAAC 372

RESULT 37
AQ362915
LOCUS   AQ362915   708 bp   DNA   linear   GSS 16-DEC-1999
DEFINITION nbxb0051A06f CUGI Rice BAC Library Oryza sativa genomic clone
            nbxb0051A06f, DNA sequence.
ACCESSION   AQ362915
VERSION   AQ362915.2 GI:6582315
KEYWORDS   GSS.
SOURCE   Oryza sativa.

```
ORGANISM    Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 708)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     On Dec 15, 1999 this sequence version replaced gi:4212570.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 118.
FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbxb0051A06f"
                     /clone_lib="CUGI Rice BAC Library"
                     /tissue_type="Leaf"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
                     HindIII; Rice is one of two most popular grains in the
                     world. Half of the world population especially those
                     inhabiting highly populated areas of the humid tropics
                     and subtropics, rely on rice as their primary source of
                     carbohydrate. Monocotyledonous rice is a diploid plant
                     (2n=24) with a haploid genome equivalent of 431 Mbp
                     (Arumuganathan and Earle, 1991). The relatively small
                     genome of rice, three times larger than that of
                     Arabidopsis, makes it suitable for genomic studies. In
                     order to facilitate positional cloning, physical mapping
                     and genome sequencing of rice, we have constructed a BAC
                     library from Oryza sativa, Nipponbare variety. The
                     library contains 36,864 clones with an average insert size
                     of 128.5 Kb providing 10.9 haploid genome equivalents. The
                     deep coverage allows the isolation of a particular sequence
                     with a probability of 99.9 %. Two high density filters,
                     each containing 18,432 clones (doubly spotted), represent
                     the whole library for colony screening."
BASE COUNT      233 a    160 c    121 g    194 t
ORIGIN

Query Match             9.9%;  Score 18;  DB 17;  Length 708;
  Best Local Similarity   100.0%;  Pred. No. 64;
  Matches  18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     42 AAGGATTTTTATTACTTA 59
          ||||||||||||||||||
Db    576 AAGGATTTTTATTACTTA 593

RESULT 38
AF136410
LOCUS       AF136410                 714 bp    mRNA    linear   EST 02-JUN-1999
DEFINITION  AF136410 Homo sapiens ovary epithelium Homo sapiens cDNA clone
            N3A018, mRNA sequence.
ACCESSION   AF136410
VERSION     AF136410.1  GI:4969451
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 714)


  AUTHORS   Zhang,X.Y., Feng,J., Li,G., Qian,H.N., Li,X.P., Wu.D.C. and Hu,Y.C.
  TITLE     Differentially expressed transcripts in ovarian cancer identified
            by mRNA differential display
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Zhang XY
            Gynecologic Oncology Center
            People's Hospital, Beijing Medical University
            133 Funei Street, Western District, Beijing, 100034, China.
FEATURES             Location/Qualifiers
     source          1..714
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="N3A018"
                     /clone_lib="Homo sapiens ovary epithelium"
                     /tissue_type="ovary epithelium"
                     /note="Anchor enzyme: NlaIII~ This SAGE library is from
                     Brain, Duke glioblastoma multiforme primary tumor derived
                     from a 51 yo male. Velculescu, V. E., Zhang, L.,
                     Vogelstein, B., and Kinzler, K. W. (1995) Serial Analysis
                     of Gene Expression. Science 270, 484-487 "
BASE COUNT      220 a    144 c    130 g    220 t
ORIGIN

Query Match             9.9%;  Score 18;  DB 9;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 64;
  Matches  18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     13 TTCACACAGACGCTGGAA 30
          ||||||||||||||||||
Db     84 TTCACACAGACGCTGGAA 101

RESULT 39
BJ172818
LOCUS       BJ172818                 733 bp    mRNA    linear   EST 24-JAN-2002
DEFINITION  BJ172818 full length cDNA library, chloronemata and young
            gametophores Physcomitrella patens subsp. patens cDNA clone
            pph33k13 3', mRNA sequence.
ACCESSION   BJ172818
VERSION     BJ172818.1  GI:18340787
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens.
  ORGANISM  Physcomitrella patens subsp. patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1 (bases 1 to 733)
  AUTHORS   Fujita,T., Shin-i,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.
            , Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
            ,M.
  TITLE     Comparison of the moss Physcomitrella patens genome with flowering
            plants genome
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            A backbone of the vector is basically from pBluescript(KS), that
            was in vivo excised from a modified lPS phage vector (Mo bi Tec,
            Germany). 5' end of the cDNA that was digested with XhoI was
            ligated to SalI site of the vector and the 3' end including polyA
            tail was ligated to BamHI site of the vector. cDNA insert could be
            amplified with conventional T7 and T3 primers. This full-length
            cDNA library was generated basically according to the method
            described in The Plant J 15, 707-720 (1998) Seki M. et al.
            Protonemata were blended by the POLYTRON, and then cultivated on
            the BCDATG medium for 13-14 days under the continuous light.
FEATURES             Location/Qualifiers
     source          1..733
                     /organism="Physcomitrella patens subsp. patens"
                     /db_xref="taxon:145481"
```

```
                                                    SOURCE      Zea mays.
                              /clone="pph33k13"     ORGANISM    Zea mays
                              /clone_lib="full length cDNA library, chloronemata and          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                              young gametophores"                                              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                              /tissue_type="mixture of chloronemata and young                 clade; Panicoideae; Andropogoneae; Zea.
                              gametophores with 2 to 5 leaves"    REFERENCE  1  (bases 1 to 755)
                                                    AUTHORS     Qiu,F., Cui,F., Guo,L., Ashlock,D.A, Wen,T.J. and Schnable,P.S.
BASE COUNT    206 a    157 c    141 g    229 t      TITLE       Expressed Sequence Tags from B73 Maize Seedlings and Silks
ORIGIN                                              JOURNAL     Unpublished (2001)
                                                    COMMENT     Contact: Patrick S. Schnable
Query Match          9.9%;  Score 18;  DB 13;  Length 733;                  Schnable Laboratory
Best Local Similarity  100.0%;  Pred. No. 64;                               Iowa State University
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;        G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
                                                                            Tel: 515-294-0975
Qy    115 GTTGATGCTGAGAACTTC 132                                            Fax: 515-294-2299
          |||||||||||||||||||                                              Email: schnable@iastate.edu
Db    288 GTTGATGCTGAGAACTTC 305                                           Individual basecall and confidence value were assigned using the
                                                                           Phred software.
                                                                           Individual basecall and confidence value were assigned using the
RESULT 40                                                                  (<http://depts.washington.edu/ventures/collabtr/direct/index.htm>#b
BH444542                                                                   rt). Overall sequence quality assessment and vector trimming were
LOCUS       BH444542        736 bp    DNA    linear  GSS 12-DEC-2001        conducted using the Lucy software (<http://www.tigr.org/softlab/>).
DEFINITION  BOGIV33TR BOGI Brassica oleracea genomic clone BOGIV33, DNA     Lucy parameters were set to ensure an overall trimmed quality of
            sequence.                                                      97.5% or better without any vector fragments in the chosen
ACCESSION   BH444542                                                       high-quality region of each sequence. Low-quality bases between the
VERSION     BH444542.1  GI:17630256                                        poly-T and the high-quality region were replaced with N's to serve
KEYWORDS    GSS.                                                           as spacers.
SOURCE      Brassica oleracea.                                             PCR PRimers
ORGANISM    Brassica oleracea                                              FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  CTA TAG)
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;   BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.     TAA AG)
REFERENCE   1  (bases 1 to 736)                                            Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
AUTHORS     Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.  FEATURES             Location/Qualifiers
TITLE       Whole genome shotgun sequencing of Brassica oleracea           source          1..755
JOURNAL     Unpublished (2001)                                                             /organism="Zea mays"
COMMENT     Other GSSs: BOGIV33TF                                                          /cultivar="B73"
            Contact: Chris Town                                                            /db_xref="taxon:4577"
            TIGR                                                                           /clone="MEST56-C10"
            9712 Medical Center Drive, Rockville, MD 20850, USA.                           /clone_lib="ISUM4-TN"
            Tel: 301-838-3523                                                              /tissue_type="Seedling and silk"
            Fax: 301-838-0208                                                             /lab_host="DH10B"
            Email: cdtown@tigr.org                                                         /note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
            DNA is from a doubled haploid provided by Tom Osborn.                          ds-cDNA molecules were generated as follows. First-strand
            Seq primer: TR                                                                 cDNA was prepared from oligo-dT selected mRNA by priming
            Class: sheared ends.                                                           with a NotI oligo-dT primer (5,
FEATURES             Location/Qualifiers                                                   AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT). The
    source          1..736                                                                 resulting DNA:RNA hybrid was treated with RNase H and used
                    /organism="Brassica oleracea"                                          as a template for DNA Pol1-catalyzed second strand
                    /strain="TO1000DH3"                                                    synthesis. After the addition of EcoRI adaptors, the
                    /db_xref="taxon:3712"                                                  ds-cDNAs were digested with NotI and size-selected. The
                    /clone="BOGIV33"                                                       resulting molecules were directionally cloned into the
                    /clone_lib="BOGI"                                                      EcoRI and NotI sites of the pT7T3PAC vector. The library
                    /note="Vector: pHOS1; Site_1; BstXI; 2-3 kb sheared                    then went through one round of normalization to CoT value
                    genomic DNA inserted into pHOS1 using BstXI linkers"                   of 5 based on the methods of Marcelo Bento Soares (Genome
BASE COUNT    268 a    149 c    170 g    149 t                                             Research 6: 791-806, 1996)."
ORIGIN                                              BASE COUNT    171 a    211 c    202 g    171 t
                                                    ORIGIN
Query Match          9.9%;  Score 18;  DB 17;  Length 736;
Best Local Similarity  100.0%;  Pred. No. 64;       Query Match          9.9%;  Score 18;  DB 13;  Length 755;
Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;  Best Local Similarity  100.0%;  Pred. No. 64;
                                                    Matches  18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
Qy    112 AAAGTTGATGCTGAGAAC 129
          |||||||||||||||||                         Qy     23 CGCTGGAAGACGTCTTCC 40
Db    656 AAAGTTGATGCTGAGAAC 673                               ||||||||||||||||||
                                                    Db    591 CGCTGGAAGACGTCTTCC 608
RESULT 41
BM072946                                            RESULT 42
LOCUS       BM072946        755 bp    mRNA    linear  EST 13-NOV-2001  BH571181
DEFINITION  MEST56-C10.T3 ISUM4-TN Zea mays cDNA clone MEST56-C10 3', mRNA   LOCUS       BH571181        763 bp    DNA    linear  GSS 14-DEC-2001
            sequence.                                               DEFINITION  BOHKK72TF BOHK Brassica oleracea genomic clone BOHKK72, DNA
ACCESSION   BM072946                                                            sequence.
VERSION     BM072946.1  GI:16916686                                 ACCESSION   BH571181
KEYWORDS    EST.
```

```
VERSION      BH571181.1  GI:17823020
KEYWORDS     GSS.
SOURCE       Brassica oleracea.
  ORGANISM   Brassica oleracea
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE    1  (bases 1 to 763)
  AUTHORS    Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE      Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL    Unpublished (2001)
COMMENT      Other GSSs: BOHKK72TR
             Contact: Chris Town
             TIGR
             9712 Medical Center Drive, Rockville, MD 20850, USA.
             Tel: 301-838-3523
             Fax: 301-838-0208
             Email: cdtown@tigr.org
             DNA is from a doubled haploid provided by Tom Osborn.
             Seq primer: TF
             Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..763
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHKK72"
                     /clone_lib="BOHK"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT     282 a    139 c    131 g    211 t
ORIGIN

Query Match      9.9%;  Score 18;  DB 17;  Length 763;
Best Local Similarity   100.0%;  Pred. No. 64;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    161 TGATTGGAATGTTCTCTT 178
          ||||||||||||||||||
Db    728 TGATTGGAATGTTCTCTT 745


RESULT 43
BI656335
LOCUS        BI656335                770 bp    mRNA      linear      EST 12-SEP-2001
DEFINITION   603282642F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5326811 5',
             mRNA sequence.
ACCESSION    BI656335
VERSION      BI656335.1  GI:15570571
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 770)
  AUTHORS    Lothar Hennighausen Ph.D., Priscilla Furth
             Ph.D.
  TITLE      cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  JOURNAL    NIH-MGC http://mgc.nci.nih.gov/,
             National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
             Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM1829 row: d  column: 12
             High quality sequence stop: 703.
FEATURES             Location/Qualifiers
     source          1..770
                     /organism="Mus musculus"
```

```
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5326811"
                     /clone_lib="NCI_CGAP_Mam4"
                     /tissue_type="tumor_gross tissue"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and Differentiation 7, 3-11 (1996)."
BASE COUNT     177 a    205 c    281 g    107 t
ORIGIN

Query Match      9.9%;  Score 18;  DB 13;  Length 770;
Best Local Similarity   100.0%;  Pred. No. 64;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     83 CAACAGCTGAGCAAGAGG 100
          |||||||||||||||||||
Db    640 CAACAGCTGAGCAAGAGG 657


RESULT 44
BH738785/c
LOCUS        BOMPT62TR  BO_2_3_KB Brassica oleracea genomic clone BOMPT62, DNA
DEFINITION                          794 bp    DNA      linear    GSS 20-FEB-2002
             sequence.
ACCESSION    BH738785
VERSION      BH738785.1  GI:18844180
KEYWORDS     GSS.
SOURCE       Brassica oleracea.
  ORGANISM   Brassica oleracea
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE    1  (bases 1 to 794)
  AUTHORS    Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE      Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL    Unpublished (2001)
COMMENT      Other GSSs: BOMPT62TF
             Contact: Chris Town
             TIGR
             9712 Medical Center Drive, Rockville, MD 20850, USA.
             Tel: 301-838-3523
             Fax: 301-838-0208
             Email: cdtown@tigr.org
             DNA is from a doubled haploid provided by Tom Osborn.
             Seq primer: TR
             Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..794
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMPT62"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT     272 a    134 c    110 g    278 t
ORIGIN

Query Match      9.9%;  Score 18;  DB 17;  Length 794;
Best Local Similarity   100.0%;  Pred. No. 65;
Matches   18;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     42 AAGGATTTTATTACTTA 59
          |||||||||||||||||
Db    546 AAGGATTTTATTACTTA 529
```

```
RESULT 45
CNS03ZRF
LOCUS           CNS03ZRF                1012 bp    DNA        linear   GSS 18-MAY-2000
DEFINITION      Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
                071E07 of library G from Tetraodon nigroviridis, genomic survey
                sequence.
ACCESSION       AL267972
VERSION         AL267972.1  GI:7989802
KEYWORDS        GSS; genome survey sequence.
SOURCE          Tetraodon nigroviridis
ORGANISM        Tetraodon nigroviridis
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
                Tetraodontidae; Tetraodon.
REFERENCE       1  (bases 1 to 1012)
AUTHORS         Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
                Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
                Saurin,W. and Weissenbach,J.
TITLE           Human gene number estimate provided by genome wide analysis using
                Tetraodon nigroviridis DNA sequence
JOURNAL         Unpublished
REFERENCE       2  (bases 1 to 1012)
AUTHORS         Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
                Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
                Weissenbach,J.
TITLE           Charaterization and repeat analysis of the compact genome of the
                freshwater pufferfish Tetraodon nigroviridis
JOURNAL         Unpublished
REFERENCE       3  (bases 1 to 1012)
AUTHORS         Genoscope.
TITLE           Direct Submission
JOURNAL         Submitted (12-APR-2000)
COMMENT         This sequence is a single read and was generated as part of a large
                scale clone-end sequencing project of the Tetraodon nigroviridis
                genome. For more information, please take a look at
                http://www.genoscope.cns.fr/Tetraodon.
FEATURES        Location/Qualifiers
     source     1..1012
                /organism="Tetraodon nigroviridis"
                /db_xref="taxon:99883"
                /clone="071E07"
                /clone_lib="G"
                /note="Genoscope sequence ID : COBG071AC04SP1-end :
                PUC-Ori"
BASE COUNT       280 a      181 c      202 g      346 t        3 others
ORIGIN
```

```
Query Match           9.9%;   Score 18;  DB 17;  Length 1012;
Best Local Similarity  100.0%;   Pred. No. 67;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     78 GAACACAACAGCTGAGCA 95
          |||||||||||||||||
Db    312 GAACACAACAGCTGAGCA 329
```

Search completed: June 9, 2003, 12:46:04
Job time : 1117 secs